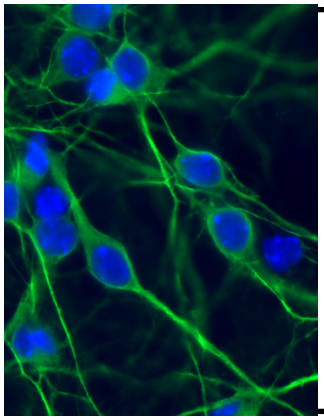
UNIVERSITE DE STRASBOURG

Institut de Génétique et de Biologie Moléculaire et Cellulaire
Illkirch-Graffenstaden

Le 15 décembre 2016

Habilitation à diriger des recherches



Etudes des bases moléculaires à l'origine du syndrome FXTAS et de la SLA/DFT.

Dr Chantal SELLIER

Membres de la commission d'examen :

Mr le Docteur Charlet-Berguerand Nicolas	Directeur de recherche
Mr le Docteur Pradat Pierre-François	Rapporteur
Mr le Docteur Sergeant Nicolas	Rapporteur
Monsieur le Docteur Dupuis Luc	Rapporteur
Madame le Docteur Martinat Cécile	Examinateur
Mr le Professeur Anheim Mathieu	Examinateur
Monsieur le Professeur Mandel Jean-Louis	Examinateur

REMERCIEMENTS

*Je tiens tout d'abord à remercier chaleureusement le **Dr Pradat Jean-François**, le **Dr Dupuis Luc**, le **Dr Sergeant Nicolas**, le **Dr Martinat Cécile**, le professeur **Mandel Jean-Louis** et le professeur **Anheim Mathieu** d'avoir accepté d'examiner mon travail.*

*Je voudrais remercier tout particulièrement **Nico** pour la confiance dont tu as fait preuve à mon égard dès mon arrivée dans l'équipe et pour tous ce que j'ai pu apprendre à tes côtés. Merci de m'avoir donné la possibilité de travailler sur les maladies neurodégénératives alors que je n'avais aucune connaissance dans ce domaine. Merci également pour ta disponibilité et toutes les discussions scientifiques que nous pouvons avoir.*

*Je veux également remercier toutes celles et ceux qui ont participé de près ou de loin à mes différentes recherches, et particulièrement tous mes co-auteurs sans qui les projets n'auraient pu aboutir. Merci également à toutes les personnes qui m'ont permis de développer de nouveaux projets, qui m'ont éclairé de leurs connaissances dans leurs domaines d'expertises et avec qui j'ai eu beaucoup de plaisir à travailler notamment l'équipe de **Cécile M.**, de **Luc D.**, et de **Frédérique G.**.*

*Merci aux différentes personnes de l'équipe Charlet : **Michel N.**, **Manon B.** et **Véronique P.** pour leur gentillesse et leur bonne humeur. Merci tout particulier à **Angéline G.** pour ton aide si précieuse dans mes différents projets depuis ton arrivée dans l'équipe et **Camille C.** qui commence l'aventure. Merci également à **Fernande F.** pour tout ce que nous avons pu partager pendant tes 5 années passées au sein de l'équipe et pour ton amitié qui est toujours présente aujourd'hui malgré la distance.*

*Un grand merci à toutes les personnes des différentes plateformes et services de l'IGBMC que j'ai côtoyées au cours de ces années et dont l'aide est si précieuse au quotidien. Merci tout particulièrement à **Alex V.** et **William M.** d'être toujours disponible dès que j'ai besoin de vous et merci également pour votre bonne humeur.*

*Merci également à tous mes collègues, à ceux qui partagent mon quotidien notamment l'**équipe Séraphin** sans qui l'ambiance serait bien triste, **Pascal E.** et **Cathy H.** qui ont toujours les mots pour rire, **Céline Z., James S.** et **Jean-Marie G.** pour leur gentillesse.*

Enfin, merci à tous ceux que je n'ai pas cités mais qui ont fait partie de cette aventure et que j'ai pu connaître tout au long de ces années.

Pour finir, merci à ma famille de m'avoir toujours soutenue et d'être toujours présente.

SOMMAIRE

PARTIE I : Curriculum vitae

Dr. Chantal SELLIER
Née le 15 octobre 1979
Nationalité française

Coordonnées professionnelles : IGBMC
Dpt de neurogénétique et médecine translationnelle
Equipe Maladies à gain de fonction d'ARN
1, rue Laurent Fries
67400 Illkirch-Graffenstaden
Tél : 03-88-65-33-57
Mail : sellier@igbmc.fr

I- SITUATION ACTUELLE

Je suis actuellement en stage post-doctoral à l'IGBMC (Institut de Génétique et de Biologie Cellulaire et Moléculaire, Illkirch-Graffenstaden) dans l'équipe "maladies à gain de fonction d'ARN" dirigée par le Dr Charlet-Berguerand (DR2 – INSERM).

II- ACTIVITES DE RECHERCHE

2003-2006 : Laboratoire de Régulation des Signaux de Division – Université de Lille 1.
Responsable : Pr Vilain Jean-Pierre.
Mon travail de recherche a porté sur l'étude de l'effet de variations du pH intracellulaire sur les voies de signalisation régulant la reprise méiotique dans les ovocytes de xénope, plus particulièrement sur l'activation de la voie MAPK/Erk2 et du MPF (mitosis promoting factor).

2007- à aujourd'hui : Equipe « Maladies à gain de fonction d'ARN »
IGBMC - Illkirch-Graffenstaden.
Responsable : Dr Charlet-Berguerand Nicolas.
Mes activités de recherche portent sur l'étude des mécanismes moléculaires impliqués dans les maladies dues à des expansions de répétitions dans des régions non codantes du génome. Je me suis focalisée plus particulièrement sur l'étude du syndrome d'ataxie et de tremblement lié à l'X fragile (FXTAS) et sur la sclérose latérale amyotrophique associée à une démence frontotemporale (SLA/DFT).

III- FORMATION

Doctorat Sciences de la Vie et de la Santé **2006**

Mention très honorable avec félicitations du jury
Directeur : Pr Vilain - Laboratoire de Régulation des Signaux de Division
Université des Sciences et Technologies de Lille 1 - Villeneuve d'Ascq

Master Biologie Cellulaire et Physiologie filière biologie cellulaire **2003**

Laboratoire de Régulation des Signaux de Division - Equipe du Pr Vilain JP
Université des Sciences et Technologies de Lille 1 - Villeneuve d'Ascq

Licence de Biologie Cellulaire et Physiologie **2001**

Université Jules Verne – Amiens

IV- ENCADREMENT D'ETUDIANTS

Encadrement de doctorants

Corbier Camille

Ecole doctorale Sciences de la vie et de la santé - Strasbourg
Septembre 2016 - à aujourd'hui
Etude du rôle de la phosphorylation de la protéine SMCR8 dans l'autophagie et de son implication dans la SLA/DFT.

Freyermuth Fernande

Ecole doctorale Sciences de la vie et de la santé - Strasbourg
Septembre 2009 - septembre 2013
Etude de l'interaction *in vitro* du complexe Drosha/DGCR8 avec les répétitions CGG dans le syndrome FXTAS (Sellier *et al.,* Cell reports, 2013).
Etude de l'interaction *in vitro* des protéines MBNL1 et rbFOX1 aux répétitions CCUG impliquées dans la dystrophie myotonique de type 2 (Sellier *et al.,* en préparation)

Encadrement d'étudiants en master 1 et 2

Boivin Manon

Master 2 Biologie et génétique moléculaire, Université de Strasbourg
Septembre 2016 - à aujourd'hui
Rôle de C9ORF72 dans la mitophagie.

Corbier Camille

Master 2 Phsyiopathologie, Université de Strasbourg
Septembre 2015 - juin 2016
Etude du rôle de la phosphorylation de la protéine SMCR8 dans l'autophagie et de son implication dans la SLA/DFT (Sellier *et al.,* 2016).

Vurpillot Séréna

Master Biotechnologies, Université des Sciences et Technologies Lille 1.
1er avril 2016 - 31 août 2016
Etude de l'interaction de l'actinine 4 avec la protéine FMRpolyGlycine dans le syndrome FXTAS.

Bruestch Anais

Master 2 Biologie du dévelopement, Université de Strasbourg
Septembre 2014 - juin 2015
Analyse du phénotype des souris knock-out pour le gène *C9ORF72*

Kreider Julie

Master 1 Neurosciences, Université de Strasbourg.
Juin - juillet 2012
Quantification des micro-ARNs altérés dans des cellules neuronales après déplétion du complexe Drosha/DGCR8

Encadrement d'ingénieurs :

Weber Pauline

5ième année école Ingénieur INSA – Lyon
Février - Août 2012
Etude de la compétition entre les protéines rbFOX1 et MBNL1 pour la liaison aux répétitions CCUG dans les dystrophies myotoniques de type 2 (Sellier *et al.,* en préparation).

Drouot Nathalie

Ingénieur d'étude
Février 2014 - juillet 2014
Etude des conséquences de la traduction des répétitions CGG dans le syndrome FXTAS

Encadrement d'étudiants en licence :

Gaucherot Angéline

Licence professionnel – Colmar
Septembre 2012 - août 2013
Développement de cellules iPS différenciées en neurones à partir de fibroblastes de patients FXTAS (Sellier *et al.,* en révision Neuron)

Seiffert Cécile

Licence biochimie et biologie moléculaire - Université de Strasbourg.
Juillet - août 2008
Etude de la localisation des différents mutants de SAM68

Encadrement d'étudiant en BTS :

Aubry Adeline

BTS Biotechnologies - Lycée Jean Rostand – Strasbourg
Mai - juin 2008
Clonage des protéines DROSHA et DGCR8 et étude de leur localisation dans les cellules neuronales.

V- ENSEIGNEMENTS

août 2003	tutorat de révision	licence 1ère et 2ième année	30 h	biologie animale, biologie cellulaire et génétique
février 2004 - avril 2004	vacations	licence 1ère année	75 h	TP embryologie
août 2003	tutorat de révision	licence 1ère et 2ième année	30 h	biologie animale, biologie cellulaire et génétique
février 2005 - avril 2005	vacations	licence 1ère année	24 h	TP sur l'étude du potentiel complexe dans le nerf sciatique de grenouille
		licence 1ère année	30 h	TP embryologie
février 2006 - avril 2006	demi-poste d'ATER	licence 3ième année	18 h	TD sur la régulation du cycle cellulaire et le cancer
		licence 2ième année	10 h	TD sur la génétique et les maladies mitochondriales
		licence 2ième année	3 h	TD portant sur la rédaction d'un mémoire thématique
		master 1ère année	25 h	TP de biologie cellulaire
		licence 3ième année	12 h	TP sur l'urogenèse chez les vertébrés
		licence 2ième année	16 h	TP sur les pièces buccales chez les insectes

VI- PUBLICATIONS

Article en préparation :

- **Sellier C***, Cerro-Herreros E, Freyermuth F, Blatter M, Hammer C, Weber P, Koch M, Ruffenach F, Page A, Sarkar P, Puymirat J**,** Udd B, Day JW, Meola G, Bassez G, Schoser B, Fujimura H, Takahashi MP, Furling D, Artero R, Allain F, Llamusi B, Charlet-B N*
  rbFOX1 chases MBNL1 from CCUG RNA aggregates and alleviates symptoms of myotonic dystrophy of type 2. (*auteur co-correspondant)

Article en révision :

- **Sellier C*,** Buijsen RA, Jung L, Tropel P, Gaucherot A, Jacobs H, Meziane H, Vincent A, Champy MF, Sorg T, Wattenhofer-Donze M, Birling MC, Eberling P, Ruffenach F, Page A, Anheim M, Martinez-Cerdeno V, Tassone F, Willemsen R, Hukema RK, Viville S, Martinat C, Todd PK, Charlet-Berguerand N*
  Translation of expanded CGG repeats into the polyglycine-containing protein FMRpolyG is pathogenic in Fragile X Tremor Ataxia Syndrome
  En révision à Neuron (*auteur co-correspondant)

Articles publiés dans des journaux internationaux à comité de lecture:

- **Sellier C***, Campanari ML, Corbier C, Gaucherot A, Kolb-Cheynel I, Oulad-Abdelghani M, Ruffenach F, Page A, Ciura S, Kabashi E, Charlet-Berguerand N*
  Loss of C9ORF72 impairs autophagy and synergizes with polyQ Ataxin-2 to induce motor neuron dysfunction and cell death.
  EMBO J. 2016 Jun 15;35(12):1276-97 (*auteur co-correspondant)
- Freyermuth F, Rau F, Kokunai Y, Linke T, **Sellier C**, Nakamori M, Kino Y, Arandel L, Jollet A, Thibault C, Philipps M, Vicaire S, Jost B, Udd B, Day JW, Duboc D, Wahbi K, Matsumura T, Fujimura H, Mochizuki H, Deryckere F, Kimura T, Nukina N, Ishiura S, Lacroix V, Campan-Fournier A, Navratil V, Chautard E, Auboeuf D, Horie M, Imoto K, Lee KY, Swanson MS, Lopez de Munain A, Inada S, Itoh H, Nakazawa K, Ashihara T, Wang E, Zimmer T, Furling D, Takahashi MP, Charlet-Berguerand N.
  Slicing misregulation of SCN5A contributes to cardiac-conduction delay and heart arrhythmia in myotonic dystrophy.
  Nat Commun. 2016 Apr 11;7:11067
- Buijsen RA, **Sellier C**, Severijnen LA, Oulad-Abdelghani M, Verhagen RF, Berman RF, Charlet-Berguerand N, Willemsen R, Hukema RK.
  FMRpolyG-positive in CNS and non-CNS organs of a fragile X premutation carrier with FXTAS.
  Acta Neuropathol Commun. 2014 Nov 26;2:162.
- **Sellier C**, Hwang VJ, Dandekar R, Durbin-Johnson B, Charlet-Berguerand N, Ander BP, Sharp FR, Angkustsiri K, Simon TJ, Tassone F.
  Decresed DGCR8 expression and miRNA dysregulation in individuals with 22Q11.2 deletion syndrome.
  PLoS One. 2014 Aug 1;9(8):e103884.

- Almeida S, Gascon E, Tran H, Chou HJ, Gendron TF, Degroot S, Tapper AR, **Sellier C**, Charlet-B N, Karydas A, Seeley WW, Boxer AL, Petrucelli L, Miller BL, Gao FB.
  Modeling key pathological features of frontotemporal dementia with C9ORF72 repeat expansion in iPSC-derived human neurons.
  Acta Neuropathol. 2013 Sep;126(3):385-99.
- Todd PK, Oh SY, Krans A, He F, **Sellier C**, Frazer M, Renoux AJ, Chen KC, Scaglione KM, Basrur V, Elenitoba-Johnson K, Vonsattel JP, Louis ED, Sutton MA, Taylor JP, Mills RE, Charlet-Berguerand N, Paulson HL.
  CGG repeat-associated translation mediates neurodegeneration in fragile X tremor ataxia syndrome.
  Neuron. 2013 May 8;78(3):440-55.
- **Sellier C**, Freyermuth F, Tabet R, Tran T, He F, Ruffenach F, Alunni V, Moine H, Thibault C, Page A, Tassone F, Willemsen R, Disney MD, Hagerman PJ, Todd PK, Charlet- N.
  Sequestration of DROSHA and DGCR8 by expanded CGG RNA repeats alters microRNA processing in FXTAS.
  Cell Reports. 2013 Mar 28;3(3):869-80.
- Disney D, Liu B, Yang WY, **Sellier C**, Tran T, Charlet-Berguerand N, Childs-Disney JL
  A small molecule that targets r(CGG)(exp) and improves defects in fragile X-associated tremor ataxia syndrome.
  ACS Chem Biol. 2012 Oct 19;7(10):1711-8.
- Marin M, **Sellier C**, Paul-Antoine AF, Cailliau K, Browaeys E, Bodart JF, Vilain JP.
  Calcium dynamics during physiological acidification in Xenopus oocyte.
  J Membr Biol. 2010 Aug;236(3):233-45.
- **Sellier C**, Rau F, Liu Y, Tassone F, Hukema RK, Gattoni R, Schneider A, Richard S, Willemsen R, Elliott DJ, Hagerman PJ, Charlet-Berguerand N.
  Sam68 sequestration and partial loss of function are associated with splicing alterations in FXTAS patients
  EMBO J. 2010 Apr 7;29(7):1248-61.
- Dehennaut V, Lefebvre T, **Sellier C**, Leroy Y, Gross B, Walker S, Cacan R, Michalski JC, Vilain JP, Bodart JF.
  O-linked N-acetylglucosaminyltransferase inhibition prevents G2/M transition in Xenopus laevis oocytes.
- **Sellier C**, Bodart JF, Flament S, Baert F, Gannon J, Vilain JP.
  Intracellular acidification delays hormonal G2/M transition and inhibits G2/M transition triggered by thiophosphorylated MAPK in Xenopus oocytes.
  J Cell Biochem. 2006 May 15;98(2):287-300.
- Bodart JF, Baert FY, **Sellier C**, Duesbery NS, Flament S, Vilain JP.
  Differential roles pf p39Mos-Xp42Mpk1 cascade proteins on Raf1 phosphorylation and spindle morphogenesis in Xenopus oocytes.
  Dev Biol. 2005 Jul 15;283(2):373-83.

Revues ou éditorial :

- Corbier C, **Sellier C***
  C9ORF72 is a GDP/GTP exchange factor for RAB8 and RAB39 and regulates autophagy.
  Small GTPases. 2016 Aug 5:0 (* auteur correspondant)
- Ciura S, **Sellier C**, Campanari ML, Charlet-Berguerand N, Kabashi E
  The most prevalent genetic cause of ALS-FTD, C9ORF72 synergizes the toxicity of ATXN2 intermediate polyglutamine repeats through the autophagy
  Autophagy. 2016 Aug 2;12(8):1406-8.
- **Sellier C**, Usdin K, Pastori C, Peschansky VJ, Tassone F, Charlet-Berguerand N.
  The multiple molecular facets of fragile X-associated tremor/ataxia syndrome.
  J Neurodev Disord. 2014;6(1):23.
- **Sellier C**, Charlet-Berguerand N.
  FXTAS : size does matter !
  Cell Cycle. 2014;13(21):3319.

Chapitre de livre

- Tassone F, **Sellier C**, Charlet-B N, Todd P
  Chapter 6 : The Molecular Biology of Premutation Expanded Alleles
  FXTAS, FXPOI, and Other Premutation Disorders
  Editors: Tassone Flora & Hall Deborah
  Springer Edition

VII- Prix scientifiques

- **Congrès (Prix de la meilleure communication orale) :**
  8th international conference on unstable microsatellites and human disease - Guanacaste - Costa Rica - 17-22 janvier 2015
  Pathological translation of CGG repeats in polyglycin protein in FXTAS
  Sellier C, Jung L, Tassone F, Viville S, Hagerman PJ, Willemsen R, Anheim M, Martinat, C, Todd PK, Charlet-B N.
- **Congrès (Prix de la meilleure communication orale):**
  14 International Fragile X Conference - Orange County CA - USA - Juillet 2014
  RNA and Protein Gain of Function in FXTAS
  Sellier C, Gaucherot A, Jung L, Tassone F, Anheim M, Viville S, Hagerman PJ, Hukema R, Willemsen R, Martinat C, Todd PK, Charlet N.
- **Recommandation par la Faculté de 1000 (2013)**
  Sellier C, Freyermuth F, Tabet R, Tran T, He F, Ruffenach F, Alunni V, Moine H, Thibault C, Page A, Tassone F, Willemsen R, Disney MD, Hagerman PJ, Todd PK, Charlet-Berguerand N.
  Sequestration of DROSHA and DGCR8 by expanded CGG RNA repeats alters microRNA processing in FXTAS.
  Cell Reports. 2013 Mar 28;3(3):869-80.

- **Congrès (Prix de la meilleure communication orale) :**
  12th International Fragile X Conference - Detroit - USA - Juillet 2010
  DROSHA/DGCR8 sequestration by expanded CGG repeats alters miRNA processing alteration in FXTAS.
  Sellier C, Hagerman PJ, Willemsen R, Charlet-BN.

- **Congrès (Prix du meilleur poster):**
  5èmes Assises de Génétique Humaine et Médicale - Strasbourg - janvier 2010
  Global micro-RNA processing alteration in FXTAS.
  Sellier C, Gattoni R, Hickel P, Willemsen R, Elliott D, Hagerman P, Charlet-B N.

VIII- COMMUNICATIONS SCIENTIFIQUES

Communications orales

- **Sellier C**, Jung L, Tassone F, Viville S, Anheim M, Martinat, C, Todd PK, Charlet-B N
  · *Pathological translation of CGG repeats in polyglycin protein in FXTAS*
  · 8th international conference on unstable microsatellites and human disease - Guanacaste Costa rica - 17-22 janvier 2015

- **Sellier C**, Gaucherot A, Jung L, Tassone F, Anheim M, Viville S, Hagerman PJ, Hukema R, Willemsen R, Martinat C, Todd PK, Charlet N.
  RNA and Protein Gain of Function in FXTAS
  14 International Fragile X Conference - Orange County CA - USA - Juillet 2014

- **Sellier C**, Fugier C, Hammer C, Udd B, Schoser B, Furling D, Charlet-Berguerand N.
  FOX chase MBNL from CCUG aggregates: an explanation to the lesser severity of DM2
  International Myotonic Dystrophy Consortium 8 - Clearwater - USA - Déc 2011

- **Sellier C**, Tabet R, Tassone F, Moine H, Willemsen R, Hagerman P, Charlet-B. N.
  Sequestration of DROSHA and DGCR8 by expanded CGG-repeats leads to altered micro-RNA processing in fragile X-associated tremor/ataxia syndrome (FXTAS)
  15th International Workshop on Fragile X and Other Early-Onset Cognitive Disorders - Berlin - Septembre 2011

- **Sellier C**, Hagerman PJ, Willemsen R, Charlet-B. N.
  DROSHA/DGCR8 sequestration by CGG repeats alters miRNA processin in FXTAS.
  12th International Fragile X Conference - Detroit - USA - Juillet 2010

- **Sellier C**, Rau F, Liu Y, Tassone F, Hukema R, Willemsen R, Hagerman P, Charlet-B N
  RNA-gain-of-function disease: Sam68 sequestration by CGG repeats in FXTAS.
  7th international myotonic dystrophy meeting - Würzburg - Allemagne - Sept 2009

- **Sellier C**, Liu Y, Tassone F, Hukema R, Willemsen R, Hagerman P, Charlet-B N.
  RNA gain of function disease: sequestration of RNA-binding protein by CGG repeats in FXTAS.
  EMBO meeting “RNA and Disease” - Rome - Mai 2008

Communications écrites

- **Sellier C**, Campanari ML, Gaucherot A, Ciura S, Kabashi E, Nicolas Charlet-B N
  Role of C9ORF72 and analysis of its depletion in neuronal cells.
  10th Brain Research Conference, Chicago, USA - 15-16 octobre 2015
- **Sellier C**, Todd PK, Tassone F, Hagerman PJ, Usdin K, Nelson DL, Willemsen R, Charlet-B N
  RNA & protein gain of function : two faces to FXTAS ?
  9th Brain Research Conference - RNA Metabolism in Neurological Disease - San Diego - USA - novembre 2013
- **Sellier C,** Freyermuth F, Fugier C, Hammer C, Udd B, Puymirat J, Bassez G, Schoser B, Allain F, Bergmann A, Furling D, Charlet-B N
  FOX proteins chase MBNL1 from CCUG aggregates: an explanation to the lesser severity of DM2?
  Functional RNAs, Cell Symposia, Sitges Spain - décembre 2012
- **Sellier C**, Freyermuth F, Fugier C, Hammer C, Udd B, Puymirat J, Bassez G, Schoser B, Allain F, Bergmann A, Furling D, Charlet-Berguerand N
  FOX proteins chase MBNL1 from CCUG aggregates: an explanation to the lesser severity of DM2?
  The complex life of mRNA, EMBO/EMBL symposia, Heidelberg, Germany - oct 2012
- **Sellier C**, Tabet R, Tasone F, Moine H, Willemesen R, Hagerman P, Charlet-B N.
  Sequestration of DROSHA and DGCR8 by expanded CGG-repeats leads to altered micro-RNA processing in Fragile X-associated Tremor/Ataxia Syndrome
  RNA binding Proteins in Neurological Disease - 6th Brain Research Conference - Washington - USA. novembre 2011
- **Sellier C**, Fugier C, Puymirat J, Bassez G, Allain F, Furling D, Charlet-B.
  Recruitment of FOX proteins within CCUG aggregates explains the differences between DM1 and DM2.
  4th International Congress of Myology - Lille - France - mai 2011
- **Sellier C**, Gattoni R, Hickel P, Willemsen R, Elliott D, Hagerman P, Charlet-B N.
  Global micro-RNA processing alteration in FXTAS.
  5èmes Assises de Génétique Humaine et Médicale - Strasbourg, France - janv 2010
- **Sellier C**, Antoine AF, Flament S, Lescuyer A, Vilain JP, Bodart JF.
  Modulation of MPF and MAPK pathways activities by physiological pHi changes.
  Cell signalling world 2006 conference - Luxembourg, Luxembourg - janvier 2006

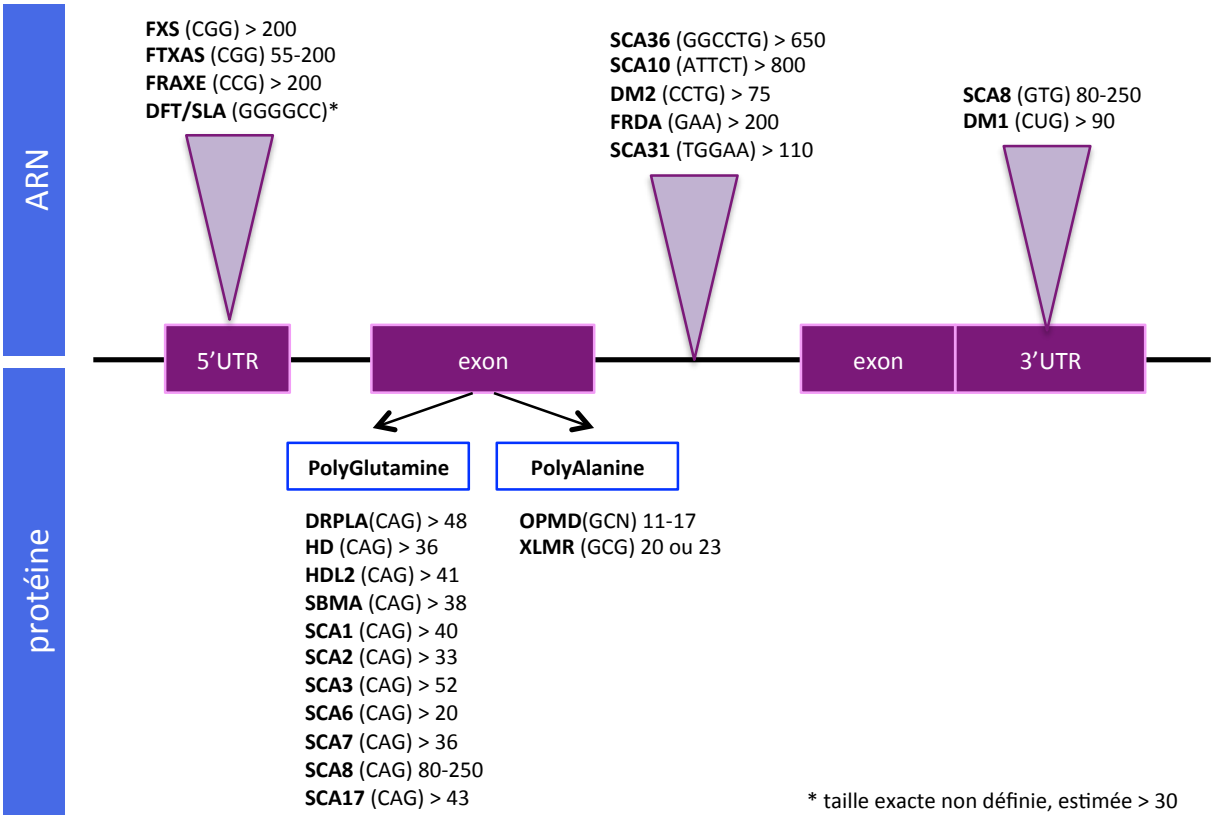
PARTIE II : Activités de recherche

I- INTRODUCTION GENERALE :

Séquences répétées dans le génome & maladies neurodégénératives

Des mutations dynamiques dues à l'instabilité de séquences répétées de nucléotides dans l'ADN sont impliquées dans une quarantaine de maladies neurodégénératives et musculaires (Loureiro *et al.*, 2016 ; figure 1). Ce type de mutation a été identifié pour la première fois en 1991 dans l'amyotrophie bulbo-spinale liée à l'X due à une expansion de répétitions CAG dans le gène du récepteur aux androgènes (Fischbeck *et al.*, 1991) et dans le syndrome de l'X Fragile dû à une expansion de répétitions CGG dans le gène *FMR1* (Fu *et al.*, 1991 ; Oberlé *et al.*, 1991 ; Pieretti *et al.*, 1991 ; Verkerk *et al.*, 1991).

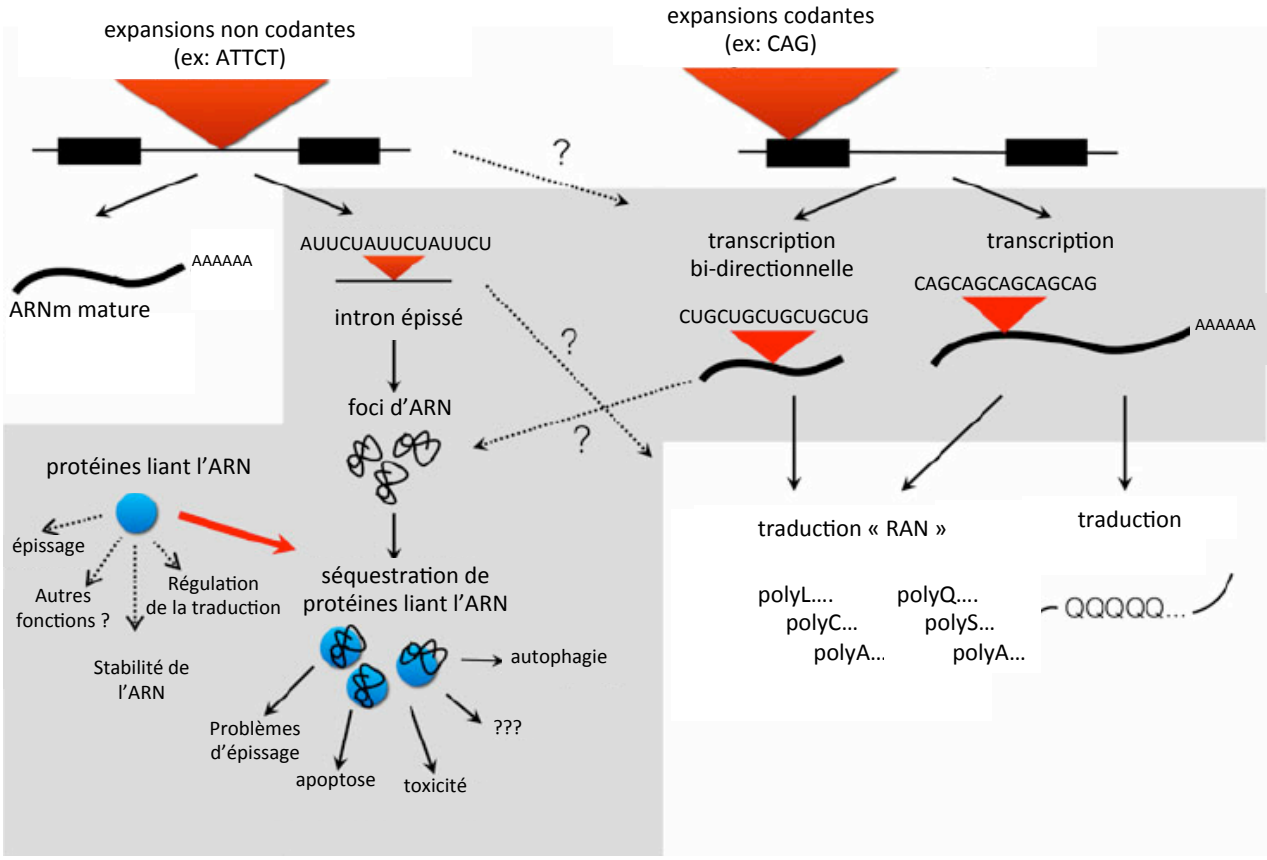
Figure 1 *: Exemples de maladies associées à des expansions de répétitions*
(modifié d'après Loureiro *et al.*, 2016)



Ces répétitions peuvent être constituées de tri-, tétra-, penta- ou hexa-nucléotides dont la longueur est polymorphique dans la population. Au-delà d'un certain seuil dit « pathologique », la maladie se développera avec plus ou moins de gravité. L'aspect « instable et dynamique » de ces répétitions fait que leur nombre peut varier chez un même individu selon le tissu, l'âge et lors de la transmission aux descendants. Il est également responsable du phénomène d'anticipation caractérisé par une diminution de l'âge d'apparition des premiers symptômes et une augmentation de la sévérité des symptômes dû à l'amplification de l'expansion au fil des générations.

Selon la localisation des répétitions dans le gène et le type de répétitions, plusieurs mécanismes distincts ont été décrits (Brouwer *et al.*, 2009 ; Loureiro *et al.*, 2016).

***Figure 2** : Schéma représentant les différents mécanismes par lesquels la présence d'expansions de répétitions est toxique dans la cellule (modifié d'après Matilla-Duenas et al., 2014)*



(1) Un gain de fonction protéique : les expansions de répétitions, essentiellement de répétitions CAG ou CAA, sont situées dans la région codante (exon) du gène résultant donc en l'expression de protéines mutées contenant des répétitions de polyglutamines (répétitions CAG) ou de polyalanines (répétitions CAA). Ces expansions de polyglutamines ou polyalanines sont pathogéniques *via* l'acquisition de nouvelles fonctions toxiques, notamment des propriétés d'agrégation, des protéines porteuses de ces répétitions. Nous pouvons ainsi citer la dystrophie musculaire oculo-pharyngée due à de courtes répétitions de codons CAA dans le gène codant pour la protéine nucléaire de liaison aux polyadénines, *PABPN1*, mais aussi l'amyotrophie bulbo-spinale liée à l'X due à une expansion de répétitions CAG traduite en polyglutamines dans le gène du récepteur aux androgènes (Fischbeck *et al.*, 1991), la maladie de Hungtington causée par des répétitions CAG dans la protéine huntingtine (Kremer *et al.*, 1994 ; Trottier *et al.*, 1995) et enfin diverses ataxies spinocérébelleuses (SCA1, SCA2, SCA3, SCA6, SCA7, etc.) dues, elles aussi, à des expansions de polyglutamines codées par des répétitions CAG dans divers gènes (respectivement dans les gènes *ATXN1*, *ATXN2, ATXN3, CACNA1A, ATXN7* pour SCA1, SCA2, SCA3, SCA6 et SCA7). Il est à noter que les expansions de CAG sont pathogènes principalement par un mécanisme de gain de fonction toxique de la protéine mutée mais aussi, par une perte partielle de la fonction normale. L'expression d'un seul allèle muté est donc suffisante pour engendrer la pathologie conduisant ainsi à une transmission autosomale dominante. L'expression de deux allèles mutés conduisant à une pathologie plus sévère ou plus précoce. De même, il est observé une corrélation entre la taille des répétitions et la sévérité et/ou l'âge d'apparition de la maladie. Des expansions plus longues ou plus exprimées (homozygotes *versus* hétérozygotes) conduisant donc à une pathologie plus sévère ou plus précoce.

(2) Une perte de fonction de la protéine codée par le gène portant les expansions. En effet, la présence d'expansion de répétitions dans des régions non codantes mais proches du promoteur du gène peut affecter la transcription de ce gène et être à l'origine de l'absence ou d'une diminution de la protéine correspondante. C'est le cas du syndrome de l'X fragile où une expansion de plus de 200 répétitions CGG dans la région 5'UTR du gène *FMR1* inhibe la transcription de ce gène et donc l'expression de la protéine FMRP. Cette protéine est essentielle pour la maturation des synapses et son absence conduit à un retard mental et un syndrome autistique (Fu *et al.*, 1991 ; Oberlé *et al.*, 1991 ; Pieretti *et al.*, 1991 ; Verkerk *et al.*, 1991). Dans le

cas du syndrome de l'X fragile, le gène *FMR1* étant situé sur le chromosome X, la transmission est évidemment liée au sexe, les hommes étant plus affectés que les femmes qui sont « protégées » par l'inactivation aléatoire d'un de leur chromosome X. Il est aussi à noter qu'il existe un effet de seuil dans ces maladies. En effet, au-delà de 200 répétitions CGG pour *FMR1*, l'expression du gène est éteinte et la sévérité de la maladie ne corrèle pas forcément avec le nombre de répétitions au contraire des maladies à gain de fonction d'ARN ou de protéine.

(3) Un gain de fonction de l'ARN : dans ce cas de figure, les expansions sont transcrites mais ne sont pas traduites et l'ARN porteur des répétitions est toxique par un mécanisme de gain de fonction. Ces répétitions sont donc typiquement situées dans des régions transcrites mais non traduites du génome (introns, 5' ou 3'UTR, etc.). Ces ARN porteurs de répétitions sont toxiques par titration de protéines liant les ARN, entraînant ainsi la perte de fonction de ces protéines séquestrées. Ainsi, dans les dystrophies myotoniques (DM), l'expression de répétitions CUG dans l'ARNm de *DMPK* (DM type 1), ou de répétitions CCUG dans le premier intron du gène *CNBP* (DM type 2) lient avec une forte affinité les facteurs d'épissage de la famille Muscleblind (MBNL1, MBNL2 et MBNL3), entraînant une perte de fonction de ces protéines, ce qui conduit à des altérations d'épissage spécifique et, *in fine,* aux symptômes de ces maladies (Miller *et al.*, 2000 ; Mankodi *et al.*, 2002 ; Lin *et al.*, 2006 ; Fugier *et al.*, 2011 ; Freyermuth *et al.*, 2016). De même, le syndrome FXTAS impliquerait un mécanisme de gain de fonction d'ARN où les répétitions CGG sont transcrites et séquestrent des protéines spécifiques (Iwahashi *et al.*, 2006 ; Sellier *et al.*, 2010, 2013). L'étude de ce mécanisme dans FXTAS a fait l'objet de la première partie de mon travail post-doctoral et sera donc abordé plus en détail dans le chapitre suivant. Tout comme les maladies à gain de fonction protéique, il est observé une transmission dominante de ces maladies à gain d'ARN, mais aussi une corrélation entre la sévérité ou l'âge d'apparition des symptômes et le nombre de répétitions.

(4) Enfin, l'équipe de Laura Ranum a décrit récemment un nouveau mécanisme conduisant à la traduction d'expansions de répétitions en absence de codon d'initiation AUG (« repeat-associated non-AUG (RAN) translation ») (Zu *et al.*, 2011). Ce travail remet donc en cause le fait que des répétitions localisées dans des régions dites non codantes ne seraient pas traduites. Ainsi, le brin antisens des répétitions CUG (donc porteur de répétitions CAG) dans la dystrophie

myotonique de type 1 ou l'ataxie spinocérébelleuse de type 8 serait traduit dans les trois cadres de lecture et donnerait des protéines porteuses de polyglutamines, de polysérines et de polyalanines (Zu *et al.*, 2011). De même, les répétitions G_4C_2 dans le gène *C9ORF72*, qui sont responsables de la sclérose latérale amyotrophique et de démence frontotemporale, codent pour des protéines porteuses de polyglycine-alanine, polyglycine-proline et polyglycine-arginine dans le brin sens, et polyproline-alanine, polyproline-glycine et polyproline-arginine dans le brin antisens (Ash *et al.*, 2013 ; Mackenzie *et al.*, 2013 ; Mori *et al.*, 2013 (a et b)). Enfin, cette découverte laisse penser que la présence de nouvelles protéines, jusqu'alors totalement insoupçonnées, pourrait contribuer à la pathogenèse de ces maladies à expansion de répétitions. Ainsi, il a été montré récemment que la traduction « RAN » de l'ARNm codant la protéine huntingtine avec des expansions de répétitions CAG conduisait à l'expression de protéines porteuses de polysérines et polyalanines (ARN sens), mais aussi polyleucines et polycystéines (ARN antisens) (Banez-Coronel *et al.*, 2015). Il semblerait donc que la traduction RAN ne soit pas limitée aux régions non codantes.

Durant mon stage post-doctoral, j'ai essentiellement axé mes recherches sur deux syndromes neurodégénératifs dus à des expansions de répétitions : le syndrome FXTAS dû à des expansions de répétitions CGG dans le gène *FMR1* et la sclérose latérale amyotrophique associée à une démence frontotemporale (SLA/DFT) due à des répétitions G_4C_2 dans le gène *C9ORF72*.

II- LE SYNDROME FXTAS

II-1- Introduction

II-1-1- Généralités

Le syndrome d'ataxie et de tremblement lié au chromosome X (FXTAS) est dû à une expansion limitée de répétitions CGG dans la région 5'UTR du gène *FMR1*. Au contraire du syndrome de l'X fragile où les expansions excèdent 200 répétitions CGG, les patients FXTAS présentent un nombre de répétitions compris entre 55 et 200 CGG, aussi connu sous le terme de prémutation. Les individus porteurs de la prémutation ont longtemps été considérés comme asymptomatiques jusqu'à l'observation par le Professeur Randy Hagerman de plusieurs cas d'ataxie dans les familles, et notamment les grands-pères, d'enfants atteints du syndrome de l'X fragile. Ce travail clinique fondateur a permis la description d'un nouveau syndrome neurodégénératif chez les porteurs d'une prémutation CGG dans le gène *FMR1* (Hagerman and Hagerman, 2001).

Les porteurs de prémutations sont estimés à 1 pour environ 250 femmes et 1 pour environ 800 hommes (Hagerman, 2008). Toutefois, le nombre de patients FXTAS est beaucoup plus rare avec une estimation de 1 individu atteint de FXTAS pour 4000 porteurs de prémutation (Jacquemont *et al.*, 2004 (a)). Ce paradoxe est dû à une pénétrance incomplète de la maladie, mais aussi par le fait que les cas de FXTAS avec des expansions en dessous de 70 – 75 répétitions CGG sont extrêmement rares (Tassone and Hagerman, 2012). En effet, on considère généralement que l'intervalle « pathologique » pour les répétitions CGG se situe environ entre 70 et 150 répétitions, alors que la plupart des porteurs de prémutation présentent moins de 70 répétitions.

II-1-2- Caractéristiques cliniques

Les deux caractéristiques majeures du syndrome FXTAS sont une démarche ataxique et/ou un tremblement d'intention qui débutent vers l'âge de 60-65 ans (Hagerman *et al.*, 2001 ; Leehey *et al.*, 2007). Les patients FXTAS peuvent également présenter des problèmes neuropsychiatriques (anxiété, agitation, dépression) et cognitifs, des symptômes parkinsoniens ainsi que des neuropathies périphériques (Bacalman *et al.*, 2006 ; Hessl *et al.*, 2005 ; Jacquemont *et al.*, 2003).

La progression de la maladie est variable avec une espérance de vie entre 5 et 25 ans (Kamm *et al.*, 2005 ; Greco *et al.,* 2006). Du fait de l'inactivation aléatoire du chromosome X, les femmes sont plus rarement atteintes et avec des symptômes plus légers que les hommes (Hagerman *et al.*, 2004).

II-1-3- Caractéristiques histopathologiques

La principale caractéristique radiologique est une atrophie généralisée du cerveau touchant plus particulièrement le cervelet mais aussi le pons et les régions frontales et pariétales du cortex (figure 3). L'IRM cérébral des patients montre une hyperintensité en T2 des pédoncules cérébelleux moyens et de la matière blanche adjacente illustrant une maladie de la substance blanche (Brunber *et al.*, 2002 ; Jacquemont *et al.*, 2003 ; Loesch *et al.*, 2005 ; Cohen *et al.*, 2006). Le degré d'atrophie du cerveau est souvent corrélé à la présence et à la sévérité du tremblement et de l'ataxie et à la taille des. Il est à noter que des études récentes montrent des changements asymptomatiques de la structure du cerveau et de la matière blanche associés à la pathologie FXTAS avant l'apparition des symptômes (Wang *et al.*, 2012 ; Battistella *et al.*, 2013).

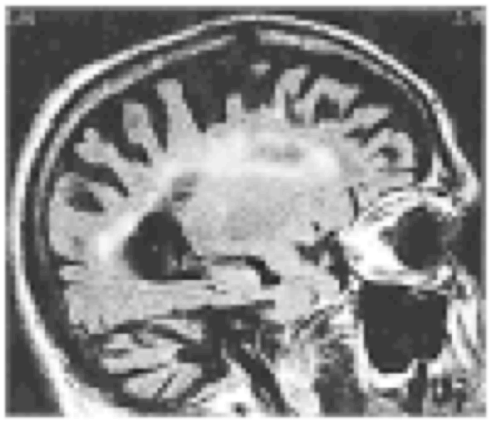


***Figure 3** : IRM montrant une sevère atrophie du cerveau chez un patient atteint de FXTAS (Jacquemont et al., 2004 (b))*

Les analyses neuropathologiques de cerveaux de patients FXTAS post-mortem montrent une perte des cellules de Purkinje, une spongiose de la substance blanche profonde du cervelet et une gliose de Bergman. De plus, le syndrome FXTAS est caractérisé par la présence d'inclusions intranucléaires éosinophiliques et positives pour l'ubiquitine dans les neurones et les astrocytes (figure 4 ; Greco *et al.*, 2002 ; Greco *et al.*, 2006). Ces inclusions contiennent également l'ARNm *FMR1* portant la prémutation (Tassone *et al.*, 2004).

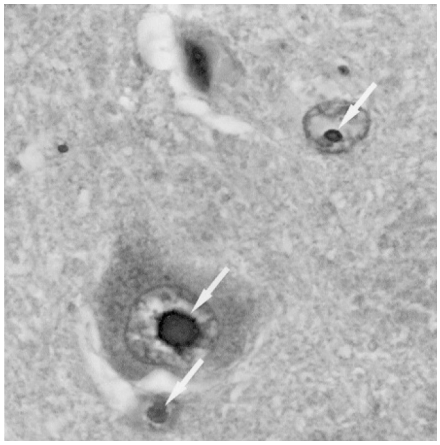


Figure 4 *: Inclusions positives à l'ubiquitine présentes dans l'hippocampe chez un patient FXTAS (Hagerman et al., 2004)*

Ces inclusions intranucléaires sont également observées dans le système nerveux périphérique incluant les cellules ganglionnaires de la medullosurrénale, les ganglions mésentériques de l'estomac et les ganglions autonomes épicardiaux du cœur (Gokden *et al.*, 2009). Plus récemment, la présence de rares inclusions intranucléaires a été notée dans des organes somatiques et neuroendocriniens (Hunsaker *et al.*, 2011). La composition de ces agrégats intranucléaires n'a été élucidée que très récemment et a occupé une grande partie de mon travail (Iwahashi *et al.*, 2006 ; Todd *et al.*, 2013 ; Sellier *et al.*, 2010, 2013, 2016 en révision).

II-1-4- Caractéristiques moléculaires

Les porteurs de prémutation présentent une augmentation de 2 à 8 fois de la quantité de l'ARNm de *FMR1* dans les lymphocytes et/ou le cerveau comparé à des échantillons contrôles (Tassone *et al.*, 2000 ; Kenneson *et al.*, 2001). Cette augmentation serait due à une transcription du gène *FMR1* plus élevée (Tassone *et al.*, 2007). Au contraire, les échantillons de patients FXTAS présentent une légère diminution de la quantité de protéine FMRP (Tassone *et al.*, 2000 ; Kenneson *et al.*, 2001). Ce paradoxe (ARN *FMR1* élevé mais protéine FMRP légèrement diminuée) est probablement dû à la difficulté du ribosome à passer à travers les fortes structures secondaires d'ARN induites par les répétitions CGG (Feng *et al.*, 1995; Primerano *et al.*, 2002).

II-1-5- Physiopathologie

Une description des mécanismes pathologiques impliqués dans FXTAS est présentée dans ce manuscrit sous forme de revue et de chapitre de livre auxquels j'ai contribué (revue 1 : Sellier *et al.*, 2014 ; Chapitre de livre : "FXTAS, FXPOI, and Other Premutation Disorders" Editeurs: Tassone Flora & Hall Deborah). Toutefois, j'aimerais décrire ci-dessous brièvement les deux mécanismes principaux (gain de fonction ARN et/ou traduction non canonique des répétitions CGG) sur lesquels j'ai travaillé.

L'un des premiers mécanismes décrit comme étant impliqué dans la pathogenèse du syndrome FXTAS est un gain de fonction toxique de l'ARNm *FMR1* présentant une expansion de répétitions CGG (Hagerman and Hagerman, 2004). En effet, l'observation que l'ARN *FMR1* porteur de répétitions CGG est présent dans les agrégats nucléaires chez les patients FXTAS (Tassone *et al.*, 2004) a conduit l'équipe du Professeur Paul Hagerman à proposer un modèle où l'ARN porteur de répétitions CGG chez les patients FXTAS puisse lier et titrer des protéines spécifiques, induisant ainsi une perte de fonction de ces protéines et donc potentiellement une mort ou un dysfonctionnement des neurones (Hagerman and Hagerman, 2004). Ce modèle est directement inspiré des dystrophies myotoniques où des ARN contenant de longues répétitions CUG titrent les facteurs d'épissage de la famille MBNL (Miller *et al.*, 2000). Diverses études sur FXTAS ont alors montré que plus d'une vingtaine de protéines étaient associées aux répétitions CGG *in vitro*, *in cellulo* ou dans les inclusions nucléaires dans des échantillons de cerveaux de patients ou des modèles drosophiles et murins de FXTAS. Ainsi, parmi ces candidats, il a été identifié des protéines de liaison à l'ARN telles que les hnRNP A1 et A2/B1, Pur-alpha, CUGBP1 ou SAM68 et DGCR8 (Iwahashi *et al.*, 2006 ; Jin *et al.*, 2007; Sofola *et al.*, 2007 ; Sellier *et al.*, 2010). Toutefois, et de façon plus surprenante, il est aussi retrouvé des protéines qui, à priori, ne devraient pas se lier à des ARN porteurs de répétitions CGG, du moins directement. Ainsi, la lamine A/C ou encore la protéine MBP sont retrouvées dans les agrégats intranucléaires dans des échantillons de cerveaux de patients FXTAS (Iwahashi *et al.*, 2006). De plus, il est noté que les agrégats intranucléaires dans FXTAS sont positifs pour l'ubiquitine et présentent une taille importante (typiquement plus de 1 à 2 µm de diamètre). Au contraire, dans les dystrophies myotoniques, l'association d'ARN porteurs de longues répétitions CUG ou CCUG et des protéines de liaison à l'ARN MBNL1 ou MBNL2 forment des agrégats nucléaires petits (0,1 à 0,2 µm de diamètre), faiblement marqués à l'ubiquitine et dont la composition ne contient pas ou très peu de protéines autres que les

protéines de liaison à l'ARN de type MBNL. Ces résultats suggèrent que les agrégats intranucléaires dans FXTAS sont clairement différents d'un simple complexe ARN-protéine.

La nature et la composition des agrégats intranucléaires caractéristiques de FXTAS et la récente découverte de la traduction des répétitions CAG en absence de codon ATG dans la dystrophie myotonique et l'ataxie spinocérébelleuse SCA8 (Zu *et al.*, 2011) ont conduit le Professeur Peter Todd et nous même à s'interroger sur une possible traduction « RAN » des répétitions CGG dans FXTAS. Il a ainsi montré que les répétitions CGG étaient effectivement traduites en polyalanines et en polyglycines dans des modèles cellulaires de FXTAS ainsi que dans des modèles de drosophile et de souris (Todd *et al.*, 2013). De plus, la présence d'agrégats de polyglycines a été confirmée indépendamment par d'autres anticorps dans des échantillons de patients FXTAS (Buijsen *et al.*, 2014). De façon intéressante, la traduction des répétitions CGG en polyalanine nécessite uniquement de longues répétitions CGG, mais la traduction en polyglycine requiert la séquence 5'UTR de *FMR1* située en aval des répétitions CGG (Kearse *et al.*, 2016). Ces résultats suggèrent des mécanismes différents de traduction des répétitions CGG, mais aussi soulèvent la question de la toxicité de ces protéines polyalanines ou polyglycines par rapport à la titration de protéines par l'ARN porteur des répétitions CGG.

II-1-6- Problématiques et approches proposées

Au début de mes recherches sur le syndrome FXTAS en 2007, le principal mécanisme proposé pour expliquer le syndrome FXTAS était un gain de fonction toxique de l'ARN *FMR1* contenant entre 55 et 200 répétitions CGG (Hagerman *et al.*, 2004). Ces ARN mutants sont proposés comme liant différentes protéines, provoquant ainsi une perte de fonction de ces dernières. Une vingtaine de protéines avaient été identifiées telles que la lamine A/C, les hnRNP A1 et A2, CUGBP1 ou Pur-alpha (Iwahashi *et al.*, 2006 ; Jin *et al.*, 2007 ; Sofola *et al.*, 2007). Je me suis donc consacrée, dans un premier temps, à l'identification de nouvelles protéines séquestrées dans les agrégats d'ARN porteurs de répétitions CGG et aux conséquences pathologiques de ces titrations (publication 1 & 2 : Sellier *et al.*, 2010 ; 2013).

Puis, la découverte en 2011 d'un nouveau mécanisme de traduction en absence de codon d'initiation ATG des expansions de répétitions (Zu *et al.*, 2011) m'a conduite à tester si les expansions de répétitions CGG dans le syndrome FXTAS pouvaient être traduites. En collaboration avec deux équipes (Peter Todd, université du Michigan, Etats-Unis; Rob Willemsen, université

Erasmus, Pays-Bas), nous avons mis en évidence que les répétitions CGG sont traduites pour donner une protéine, nommée FMRpolyGlycine, qui est toxique pour les neurones et qui forme des agrégats nucléaires positifs pour l'ubiquitine. Cette dernière est présente dans les inclusions intranucléaires caractéristiques des patients FXTAS (publication 3 & 4 : Todd *et al.*, 2013 ; Buijsen *et al.*, 2014). Je me suis donc intéressée plus particulièrement à cette protéine, notamment en étudiant par quels mécanismes elle peut-être toxique pour les neurones. De plus, nous avons généré des cellules iPS de patients FXTAS et deux nouveaux modèles murins afin d'étudier le rôle respectif des agrégats d'ARN CGG et des agrégats protéiques FMRpolyG dans la pathogenèse du syndrome FXTAS (Publication 5 : Sellier *et al.*, en révision).

II-2- Résultats

II-2-1- Identification de nouvelles protéines séquestrées dans les agrégats d'ARN (Publication 1 & 2) :

L'ARN *FMR1* contenant des expansions de répétitions CGG s'accumule dans les inclusions intranucléaires chez les patients FXTAS (Tassone *et al.*, 2004), ce qui pourrait conduire à la séquestration et à la perte de fonction de protéines spécifiques (Hagerman and Hagerman, 2004 ; Iwashi *et al.*, 2006). Nous avons donc recherché de nouvelles protéines pouvant se lier spécifiquement aux expansions des répétitions CGG. Pour cela, j'ai utilisé des ARN synthétiques composés de trente à soixante répétitions CGG couplés à des billes de sépharose que j'ai incubés avec des extraits de protéines de cerveaux de souris. Après lavage et élution des protéines liées aux répétitions CGG, nous avons identifié celles-ci par spectrométrie de masse avec l'aide de la plateforme de protéomique de l'IGBMC. J'ai pu ainsi confirmer plusieurs protéines, comme les hnRNP A1 et A2/B1 identifiées dans des études précédentes (Iwashi *et al.*, 2006), mais aussi mettre en évidence que le facteur d'épissage SAM68 (publication 1) et le complexe Drosha/DGCR8 (publication 2) sont présents spécifiquement dans les agrégats d'ARN CGG.

Le facteur d'épissage SAM68 est exprimé de manière ubiquitaire avec une expression élevée dans le système nerveux central, notamment dans le cortex, l'hippocampe, le striatum et le cervelet (Grange J, *et al.*, 2004). SAM68 régule l'épissage de différents pré-ARNm comme *Bcl-x* (Paronetto *et al.*, 2007) ou l'exon 7 de *SMN2* (Pedrotti *et al.*, 2010). Enfin, les souris knockout pour SAM68 présentent un défaut de coordination motrice (Lukong and Richard, 2008).

J'ai montré que l'expression de répétitions CGG_{60x} dans des cellules en culture aboutit à la séquestration d'environ 30% de la protéine SAM68 dans les agrégats d'ARN CGG_{60x} (Publication 1, figure 6). Nous avons également confirmé que le recrutement de SAM68 dans les agrégats d'ARNm CGG chez les patients FXTAS entraîne une altération de l'épissage de ses cibles (Publication 1, figure 8). Il apparaît donc que la protéine SAM68 est partiellement séquestrée dans les agrégats d'ARN CGG, conduisant à une diminution de son activité de régulation de l'épissage alternatif. Cependant, l'altération de la fonction de SAM68 n'est que partielle et pourrait donc ne jouer qu'un faible rôle dans la toxicité des répétitions CGG. De plus, nous avons aussi montré que la protéine SAM68 ne se lie pas directement aux expansions de répétitions CGG, mais que sa présence dans les agrégats d'ARN CGG nécessite un intermédiaire. J'ai alors identifié cet intermédiaire comme étant la protéine DGCR8 (Publication 2). Le complexe

Drosha/DGCR8 est responsable de la maturation des pri-micro-ARNs en pré-micro-ARNs (Lee *et al.*, 2003). Notre travail sur FXTAS montre que la protéine DGCR8 peut se lier directement à l'ARN composé de répétitions CGG et que cette liaison entre en compétition avec l'association de DGCR8 aux pri-micro-ARNs. En conséquence, en présence d'un excès d'ARN contenant des répétitions CGG, le complexe Drosha/DGCR8 est titré par ces répétitions et n'est donc pas disponible pour cliver les pri-micro-ARNs en pré-micro-ARNs, conduisant alors à une diminution globale du niveau des micro-ARNs chez les patients FXTAS (Publication 2, figure 4). Enfin, la toxicité cellulaire induite par l'expression de répétitions CGG peut être diminuée par l'expression concomitante de DGCR8 (Publication 2, figure 5). Nous proposons donc que la titration du complexe Drosha/DGCR8 par les répétitions CGG soit en partie responsable de la toxicité de ces répétitions.

En conclusion, mes résultats ont permis d'identifier de nouvelles protéines présentes dans les agrégats d'ARN CGG qui pourraient être impliquées dans le syndrome FXTAS. Cependant, la titration du complexe Drosha/DGCR8 et surtout de SAM68 par les répétitions CGG n'est que partielle et n'explique ni la taille, ni la présence d'ubiquitine dans les inclusions intranucléaires des patients FXTAS. Suite à la découverte d'une traduction non canonique des répétitions CAG (Zu *et al.*, 2011), nous avons alors testé la possibilité qu'un autre mécanisme, la traduction des répétitions CGG en une protéine toxique, puisse être impliqué dans la pathogenèse du syndrome FXTAS.

II-2-2- Traduction des expansions de répétitions CGG (Publications 3, 4) :

En collaboration avec l'équipe du Professeur Todd Peter, nous avons montré que les expansions de répétitions CGG pouvaient être traduites en absence de codon d'initiation ATG, et ceci dans des cultures cellulaires et dans des modèles de drosophiles (Publication 3, figures 1 et 2). Ces résultats suggèrent que le mécanisme de traduction des expansions de répétitions CAG en absence de codon d'initiation AUG (Zu *et al.*, 2011) s'applique également aux expansions de répétitions CGG dans le syndrome FXTAS (Publication 3). Toutefois, et au contraire du travail de Zu *et al.* (2011), la traduction des répétitions CGG semble se faire préférentiellement dans une seule phase codante (polyglycine) et nécessite la présence de la région 5'UTR de l'ARN *FMR1* située avant les répétitions CGG. Le produit majeur de cette traduction est une petite protéine d'environ 10 à 14 kilodaltons qui contient une expansion de glycines dont le nombre correspond

au nombre de répétitions CGG. L'expression de cette protéine, nommée FMRpolyG, est toxique pour les cellules neuronales et conduit à la formation d'agrégats intranucléaires positifs pour l'ubiquitine (publication 3). Avec l'aide du service de production d'anticorps de l'IGBMC, nous avons généré des anticorps monoclonaux contre FMRpolyG et montré que cette protéine est retrouvée spécifiquement dans les inclusions nucléaires chez les patients FXTAS dans différentes parties du cerveau comme l'hippocampe, le cortex frontal ou le cervelet mais également dans des organes n'appartenant pas au système nerveux central comme la thyroide, le cœur ou encore les reins (Publication 4). Enfin, FMRpolyG est également détectée dans des modèles murins de FXTAS (Todd *et al.*, 2013 ; Hukema *et al.*, 2015).

II-2-3- Rôle des agrégats d'ARN et de la protéine FMRpolyG dans le syndrome FXTAS (Publication 5) :

L'expression d'une protéine traduite depuis les répétitions CGG dans FXTAS soulève de nombreuses questions, notamment sur le mécanisme d'une traduction en absence de codon d'initiation ATG, mais aussi sur les différences entre la traduction principalement dans une seule phase de lecture pour les répétitions CGG contre la traduction dans les trois phases de lecture observée pour les répétitions CAG ou GGGGCC (Zu *et al.*, 2011 ; 2013). De plus, l'expression de cette protéine FMRpolyG pose la problématique de sa toxicité et notamment sa pertinence par rapport à la séquestration des protéines SAM68 ou Drosha/DGCR8 par les répétitions CGG au niveau ARN.

Dans un premier temps, un travail de construction de mutants et d'analyse protéomique m'a permis de mettre en évidence que la traduction de la protéine FMRpolyG requiert la présence d'un codon d'initiation faible ACG qui se trouve en amont des répétitions CGG. Une initiation non canonique à des codons dits d'initiation faible (ACG, CTG, GTG, etc.) est connue depuis le travail de Marilyn Kozac (Kozac 1988 ; 1990) et expliquerait ainsi la traduction des répétitions CGG principalement dans une seule phase codante (celle du codon ACG). L'expression de FMRpolyG principalement chez les patients FXTAS et non dans les échantillons contrôles résulte probablement de l'augmentation de 2 à 8 fois de l'expression de l'ARNm *FMR1* (Tassone *et al.*, 2000 ; Kenneson *et al.*, 2001), ainsi que de la présence de l'expansion de répétitions CGG qui augmente d'autant plus la taille et donc la stabilité de la protéine FMRpolyG chez les patients FXTAS (Publication 5).

Dans un deuxième temps, je me suis plus particulièrement intéressée à la toxicité de la protéine FMRpolyG. Des expériences de transfection en cultures neuronales montrent que la toxicité de cette protéine est due en partie à sa région C-terminale alors que l'expansion de polyglycine est responsable de l'agrégation de la protéine. De plus, nous avons montré que FMRpolyG interagit avec la protéine LAP2β (lamina-associated polypeptide 2ß). LAP2β est une protéine essentielle pour ancrer la lamine B à la membrane nucléaire et donc pour organiser l'architecture nucléaire. L'interaction de FMRpolyG avec LAP2β conduit à une localisation anormale de LAP2, qui est associée à une altération de la structure de la lamine nucléaire (publication 5). Il est à noter qu'une altération de la lamine nucléaire dans FXTAS était connue depuis plusieurs années mais sans qu'aucune explication ou mécanisme ne puisse expliquer cette altération (Arocena *et al.*, 2005). Nous avons alors confirmé une altération de la localisation de LAP2β et de la lamine nucléaire dans des neurones différenciés à partir d'iPS de patients FXTAS, mais aussi dans des lames de cerveaux de souris modèles de FXTAS ainsi que dans des échantillons de cerveaux de patients FXTAS (Publication 5, figure 6). Enfin, l'expression de LAP2β permet de corriger les altérations de la lamine nucléaire et de diminuer la toxicité neuronale induite par l'expression de FMRpolyG. La présence de la protéine LAP2β dans les agrégats de protéines FMRpolyG accompagnée d'une destructuration de la lamine B semble donc être un évènement pathogénique important pour comprendre FXTAS.

Enfin, dans le but de comprendre l'importance respective de l'ARN porteur des répétitions CGG comparée à la traduction de la protéine FMRpolyG pour la pathogénèse de FXTAS, nous avons généré et comparé deux modèles murins. Le premier exprime seulement les répétitions CGG au niveau ARN (CGG_{99x}). Au contraire, notre deuxième modèle de souris possède la région 5'UTR de *FMR1* contenant le codon d'initiation non canonique ACG en plus des répétitions CGG (5'UTR CGG_{99x}). Ce deuxième modèle murin exprime donc les répétitions CGG au niveau ARN mais aussi la protéine FMRpolyG. Ces modèles animaux nous ont ainsi permis d'analyser plus spécifiquement la toxicité induite par l'ARN CGG ou par la protéine FMRpolyG (Publication 5). Les résultats obtenus ont montré clairement que seule l'expression de FMRpolyG est toxique. En effet, les souris exprimant FMRpolyG présentent des altérations locomotrices dès l'âge de 3 mois. Au contraire, les souris exprimant les répétitions CGG seulement au niveau ARN ne montrent aucun symptôme dans les quinze mois de notre étude. Nous avons contrôlé que l'expression des répétitions CGG était équivalente entre les deux

modèles souris. De plus, nous avons exclu un effet toxique non spécifique de l'insertion des transgènes en choisissant une approche de transgénèse par recombinaison homologue dans le locus neutre Rosa26. Il est fortement possible que les répétitions CGG soient toxiques au niveau ARN seulement après accumulation sur une période de temps se comptant en dizaines d'années, ou que cette toxicité ne puisse pas être modélisable en modèles murins pour d'autres raisons. Toutefois et dans les limites inhérentes aux modèles animaux, nous pouvons conclure que l'expression des répétitions CGG au niveau ARN n'est que faiblement toxique, alors que l'expression de FMRpolyG induit clairement une pathologie chez l'animal.

II-3- Conclusion & discussion

En conclusion, au moins deux mécanismes pathogéniques distincts et non exclusifs seraient impliqués dans le syndrome FXTAS. Dans un premier temps, un modèle de gain de fonction toxique de l'ARN *FMR1* contenant de longues expansions de répétitions CGG a été mis en évidence, similairement à celui initialement décrit dans les dystrophies myotoniques (Lee and Cooper, 2009 ; Wheeler and Thornton, 2007). Les expansions de tailles pathologiques de répétitions CGG séquestrent différentes protéines et conduisent à une perte de fonction de celles-ci. Une trentaine de protéines liant directement ou indirectement les répétitions CGG ont été identifiées (Iwashi *et al.*, 2006, ; Jin *et al.*, 2007 ; Sofola *et al.*, 2007). Mon travail a permis d'identifier deux nouvelles protéines présentes dans les agrégats d'ARN, notamment le complexe Drosha/DGRC8 *via* une interaction directe de la protéine DGCR8 aux expansions de répétitions CGG, ce qui aboutit à une diminution de l'activité du complexe Drosha/DGRC8 et donc *in fine* à une diminution de la quantité de micro-ARN matures dans les échantillons de cerveaux de patients FXTAS (Sellier *et al.*, 2013). Par ailleurs, j'ai pu montrer que le facteur d'épissage SAM68 est partiellement séquestré dans les agrégats d'ARN CGG, mais cette liaison est indirecte et nécessite une interaction avec la protéine DGCR8 (Sellier *et al.,* 2010). Cette titration partielle de SAM68 entraîne des altérations de l'épissage alternatif des cibles de SAM68, comme *ATP11b* ou *SMN2*. Toutefois, il est à noter que ces altérations d'épissage ou d'expression de micro-ARNs sont partielles et relativement faibles, notamment si l'on compare ces altérations à celles observées dans les dystrophies myotoniques (une maladie typiquement à gain de fonction ARN). Ces observations couplées à la découverte de la traduction des répétitions CGG en protéine FMRpolyG (Todd *et al.*, 2013), nous ont conduits à remettre en question le modèle de gain de fonction ARN dans FXTAS. Nous avons alors produit des anticorps confirmant l'expression de cette protéine chez les patients et les modèles murins FXTAS (Buijsen *et al.*, 2014). De plus, nous avons aussi généré deux nouveaux modèles murins pour étudier l'impact de la seule expression de l'ARN contenant les répétitions CGG aux conséquences de l'expression de la protéine FMRpolyG. A notre surprise, seules les souris exprimant FMRpolyG montrent un phénotype. De plus, l'étude de nos modèles murins ainsi que de neurones différenciés à partir d'iPS de patients FXTAS montrent que les agrégats de protéines FMRpolyG apparaissent avant les agrégats d'ARN CGG. Ces résultats suggèrent que l'apparition des agrégats d'ARN est un événement tardif dans FXTAS alors que l'expression de la protéine FRMpolyG aurait un rôle précoce et majeur dans le

développement de la maladie. Ces résultats suggèrent ainsi un modèle principalement de gain de fonction protéique pour FXTAS.

Il est important de noter que nous n'avons pas détecté la présence de protéine issue de la traduction des CGG dans les deux autres cadres de lecture (polyalanine ou polyarginine). De plus, mes résultats montrent que la traduction de la protéine FMRpolyG est initiée en amont des répétitions CGG au niveau d'un codon d'initiation faible ACG. La présence de ce dernier est nécessaire et suffisante pour initier la traduction de la protéine FMRpolyG indépendamment de la taille de l'expansion des répétitions CGG. Ces résultats suggèrent la présence d'une uORF (upstream ORF) qui régule l'expression du gène *FMRP* (Publication 5). La présence d'une uORF dans *FMR1* n'est pas forcément inattendue sachant que des études bioinformatiques et des expériences de profilage de ribosome récentes montrent qu'environ 50% des gènes humains ou murins contiennent des uORFs et que la moitié de ces uORFs sont traduites à partir de codons d'initiations non AUG tels que CUG, GUG, ACG (Ingolia *et al.*, 2011 ; Lee *et al.*, 2012). Ces uORFs sont généralement de petite taille et ne codent que pour quelques acides aminés ou de courts peptides (<10 kDa), qui sont difficilement détectables du fait de leur petite taille et/ou de leur instabilité (Slavoff *et al.*, 2013). Le rôle principal des uORFs ne serait donc pas de coder pour des protéines fonctionnelles, mais plutôt de réguler (le plus souvent négativement) la traduction de l'ORF principale située en aval. Dans le cas de la protéine FMRpolyG, nous ne détections effectivement pas cette protéine dans les échantillons contrôles, nous pouvons supposer que la petite taille de cette protéine (<6 kDa avec 30 répétitions CGG) rend particulièrement difficile sa détection. Au contraire, l'expansion au-delà de 50 répétitions CGG chez les patients FXTAS permet d'augmenter le poids moléculaire de FMRpolyG et permet ainsi de la stabiliser (~10 à 14 kDa avec 100 ou 150 répétitions CGG). Par ailleurs, il est également probable que l'augmentation de la quantité d'ARNm *FMR1* chez les patients FXTAS ainsi que la présence des répétitions CGG qui diminuent la vitesse de déplacement du ribosome sur l'ARNm augmentent ainsi les probabilités d'initiation au codon d'initiation faible ACG et par conséquent l'expression de FMRpolyG.

Concernant une traduction éventuelle des répétitions CGG en protéine polyalanine (Todd *et al.*, 2013), nous avons généré des anticorps dirigés contre cette protéine mais nous ne l'avons ni observée dans nos modèles cellulaires ou murins de FXTAS ni dans des échantillons de patients FXTAS. Cependant, nos constructions possédant le 5'UTR du gène *FMR1*, il est possible que

l'initiation au codon non canonique ACG soit supérieure en efficacité à une traduction « RAN » initiant directement dans les trois phases codantes à l'intérieur des répétitions CGG. Ces différences ainsi que les mécanismes de la traduction RAN restent donc à étudier et à clarifier.

II-3- Projet de recherche

II-3-1- Développement de souris avec une expression spécifique dans le cervelet de la protéine FMRpolyG

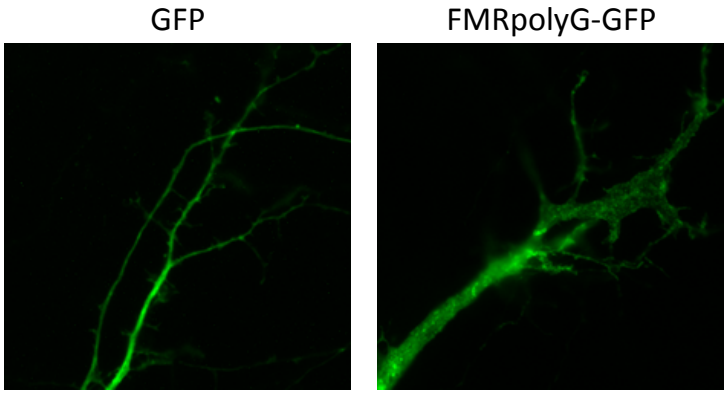
Lors de la génération de nos modèles de souris, nous avons limité l'expression de nos transgènes par une cassette loxP stop. Dans un premier temps, nous avons utilisé une recombinase CRE constitutive exprimée sous le promoteur CAG. Ce dernier permet une expression ubiquitaire et forte du transgène (Niwa *et al.*, 1991). Les souris exprimant la protéine FMRpolyG développent des problèmes locomoteurs dès l'âge de trois mois, mais aussi une obésité à partir de l'âge de 6 mois empêchant toutes analyses phénotypiques à des temps plus tardifs. Les patients FXTAS ne sont pas particulièrement décrits comme obèses et l'expression du gène *FMR1* est essentiellement restreinte au système nerveux central et aux testicules (Hinds *et al.*, 1993). Nous avons donc émis l'hypothèse que l'expression forte du transgène hors du système nerveux central puisse être responsable de l'obésité notamment en affectant le métabolisme de l'insuline ou les hormones qui régulent la satiété telles que la leptine libérée par le tissu adipeux blanc ou la ghréline, un peptide sécrété par l'estomac ou le duodénum (Leibel, 2008 ; Woods, 2009 ; Kaiyala and Schwartz, 2011). Pour éviter une contribution non neuronale au phénotype de nos souris, nous avons alors induit l'expression de FMRpolyG uniquement dans les neurones en utilisant un promoteur Nestin Cre. Cependant, et bien que nous confirmons une expression strictement neuronale de FMRpolyG dans ces nouvelles souris, ces animaux présentent eux aussi une obésité sévère à partir de 6 mois avec des souris qui atteignent plus de 70 grammes à 9 mois d'âge. Une telle obésité est typique des souris ob/ob qui sont knockout pour les récepteurs neuronaux à la leptine et donc ne présentent plus d'inhibition de leur satiété. Par RT-PCR, nous avons montré que l'expression de ce récepteur est altérée dans le cerveau des souris exprimant FMRpolyG (que ce soit sous la dépendance de la Nestin CRE ou de la CAG CRE). Ces résultats suggèrent que l'expression de FMRpolyG dans le système nerveux central est suffisante pour induire des dysfonctionnements neuronaux, notamment dans les zones qui régulent le contrôle de la satiété comme l'hypothalamus ou les régions extra-hypothalamiques. Cependant, l'obésité n'étant pas caractéristique du syndrome FXTAS, nous n'avons pas approfondi les recherches concernant ce phénotype. Pour éviter ce biais d'expression excessive de FMRpolyG dans l'hypothalamus de souris, nous développons actuellement un nouveau

modèle de souris dans lequel l'expression de FMRpolyG est contrôlée par une PCP2 CRE, limitant ainsi l'expression du transgène aux cellules de Purkinje (Nordquist *et al.*, 1988 ; Oberdick *et al.*, 1990). Nous espérons ainsi que ce nouveau modèle permette d'étudier des altérations du cervelet sur des temps plus longs.

II-3-2- Etude de l'altération de la lamina nucléaire:

Mes résultats mettent en évidence que la protéine FMRpolyG interagit avec LAP2β, entraînant une perte de localisation de cette dernière au niveau de la membrane interne du noyau. Cette altération est associée à une altération de la localisation de la lamine B et de l'architecture de la lamina nucléaire. LAP2β contient un domaine transmembranaire permettant son ancrage à la membrane interne du noyau et un domaine LEM (LAP2, Emerin, MAN1 domain), qui interagit avec les lamines B1 ou B2 (Dechat *et al.*, 2000 ; Lee *et al.*, 2001 ; Shumaker *et al.*, 2001). De ce fait, une diminution d'expression de LAP2β conduit à une altération de l'assemblage de la lamine B dans la lamina nucléaire (Furukawa and Kondo, 1998). En accord avec un rôle important de LAP2β dans la pathogénèse de FXTAS, j'ai observé que la transfection de LAP2β dans des cellules exprimant FMRpolyG corrige les altérations de la lamine B et la mort cellulaire (publication 5), suggérant que l'interaction de FMRpolyG avec LAP2β conduit à une perte de fonction de cette protéine. J'aimerais désormais étudier d'autres protéines de la lamina nucléaire, telles que les protéines ancrées à la membrane interne du noyau (Emerine, MAN1, Lap1, etc.), les lamines A et C, mais aussi les protéines SUN et Nesprine qui font le lien entre la lamina nucléaire et le cytosquelette d'actine. Ces protéines me paraissent d'autant plus intéressantes à étudier que l'expression de FMRpolyG, bien que nucléaire, conduit clairement à des altérations du cytosquelette avec la présence de dendrites anormales et élargies (figure 5).

Figure 5 : *Neurites de cultures primaires de cortex de souris après transfection de GFP ou de FMRpolyG-GFP pendant 24h (x63).*



De plus, LAP2β interagit avec la protéine BAF (barrier-to-autointegration factor ; Furukawa, 1999 ; Segura-Totten and Wilson, 2004), une protéine essentielle qui lie l'ADN double brin *in vitro* (Zheng *et al.*, 2000) et qui réprime l'activité du facteur de transcription CRX (cone-rod homeobox) *in vivo* (Wang *et al.*, 2002). De même, LAP2β lie le régulateur de transcription GCL (Germ cell less) qui est exprimé de manière ubiquitaire et localisé à proximité de la membrane interne du noyau (Jongens *et al.*, 1994 ; Nili *et al.*, 2001). Le facteur CGL lie et inhibe la sous-unité DP3 du facteur de transcription E2F-DP3. La surexpression de LAP2β est suffisante pour réprimer les gènes rapporteurs dépendant de E2F-DP3 (de la Luna *et al.*, 1999 ; Nili *et al.*, 2001). Ces études suggèrent donc que LAP2β régule la transcription *via* son interaction avec différentes protéines (BAF/ CRX, GCL/ E2F-DP3, etc.). Il serait donc intéressant d'étudier par séquençage d'ARN le transcriptome de cellules neuronales exprimant FMRpolyG. Ayant développé plusieurs modèles pour FXTAS, nous pourrons alors valider par qRT-PCR d'éventuelles altérations transcriptionelles dans nos modèles de souris mais aussi de neurones dérivés d'iPS de patients FXTAS. Ces expériences devraient nous permettre d'étudier si la présence de LAP2β dans les agrégats de protéine FMRpolyG conduit à d'éventuelles altérations de la transcription. Nous pourrons alors essayer de corriger ces altérations par l'expression de LAP2β soit par transfection en cellules neuronales immortalisées, soit par l'utilisation de vecteurs viraux de type lentivirus pour les neurones obtenus à partir d'iPS de patients ou AAV2/9 par injections dans le cerveau de souris FXTAS. Ces expériences d'injection dans la souris sont lourdes, longues et relativement complexes, toutefois, en cas de succès elles permettraient d'établir définitivement un rôle de LAP2β dans la pathogénèse de FXTAS.

REVIEW Open Access

The multiple molecular facets of fragile X-associated tremor/ataxia syndrome

Chantal Sellier[1], Karen Usdin[2], Chiara Pastori[3], Veronica J Peschansky[3], Flora Tassone[4,5] and Nicolas Charlet-Berguerand[1,6*]

Abstract

Fragile X-associated tremor/ataxia syndrome (FXTAS) is an adult-onset inherited neurodegenerative disorder characterized by intentional tremor, gait ataxia, autonomic dysfunction, and cognitive decline. FXTAS is caused by the presence of a long CGG repeat tract in the 5′ UTR of the *FMR1* gene. In contrast to Fragile X syndrome, in which the *FMR1* gene harbors over 200 CGG repeats but is transcriptionally silent, the clinical features of FXTAS arise from a toxic gain of function of the elevated levels of *FMR1* transcript containing the long CGG tract. However, how this RNA leads to neuronal cell dysfunction is unknown. Here, we discuss the latest advances in the current understanding of the possible molecular basis of FXTAS.

Review

Introduction

Fragile X-associated tremor/ataxia syndrome (FXTAS) is a neurodegenerative disorder that affects older adults who have a large CGG-repeat tract in the 5′-untranslated region (UTR) of the *Fragile X Mental Retardation 1* (*FMR1*) gene [1]. Historically, carriers of Fragile X (FX) premutation alleles with 55 to 200 CGG repeats are considered at risk for FXTAS. The prevalence of premutation alleles is approximately 1 in 260 to 1 in 800 for males and 1 in 130- to 1 in 250 for females in the general population [2,3]. Given reduced penetrance of FXTAS, it is estimated that 1 in 2,000 men over the age of 50 years in the general population will show symptoms of FXTAS [2,4-6]. Clinical features of FXTAS include progressive intention tremor and gait ataxia, which is frequently accompanied by progressive cognitive decline, parkinsonism, peripheral neuropathy, and autonomic dysfunction [7]. The neuropathology of FXTAS consists of mild brain atrophy and degeneration of the cerebellum, including hyperintensity of the middle cerebellar peduncle (MCP), loss of Purkinje neuronal cells, spongiosis of the deep cerebellar white matter, Bergman gliosis, and swollen axons [8-10]. Immunocytochemical staining of post-mortem brain tissue from FXTAS reveals the presence of eosinophilic and ubiquitin-positive intranuclear inclusions that are broadly distributed throughout the brain, including in neurons and astrocytes [9,10], the spinal column, and several non-nervous tissues including thyroid, heart, and the Leydig cells in the testes [11,12].

In contrast to the absence of *FMR1* mRNA and protein expression seen in carriers of a full mutation (over 200 CGG repeats), individuals with premutation alleles have markedly increased expression of *FMR1* mRNA, but only moderately decreased FMRP levels [13-16]. FXTAS is not seen in carriers of fully silenced *FMR1* alleles, suggesting that a novel mechanism, involving increased expression of the long CGG repeat tract in the *FMR1* mRNA, is responsible for FXTAS. In support of this hypothesis, multiple studies have demonstrated adverse consequences of expressing CGG repeats in fly, mouse, and cell models [17-22]. Consistent with an RNA-based pathological mechanism, FXTAS has also been reported in individual carriers of intermediate alleles (45 to 55 CGG repeats) [23,24], and in full mutation allele carriers who are mosaics, both for repeat size and methylation, and who still express some *FMR1* mRNA [25-27]. In addition, there has been a report documenting the presence of intranuclear inclusions in the brains of three older adult males with Fragile X syndrome (FXS) [12]. These results have implications for the spectrum of FX-associated disorders, and suggest that the definition of FXTAS may need to be

* Correspondence: ncharlet@igbmc.fr
[1]Department of Translational Medicine, IGBMC, INSERM U964 Illkirch, France
[6]Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS UMR7104, INSERM U964, University of Strasbourg, 1 rue Laurent Fries, Illkirch F-67404, France
Full list of author information is available at the end of the article

broadened to include individuals whose *FMR1* allele, irrespective of its size, makes sufficient RNA for its deleterious effects to be apparent.

How the RNA containing expanded CGG repeats leads to FXTAS pathogenesis is not yet fully known. This review will cover the recent advances in the understanding of the molecular mechanisms that may contribute to the pathogenesis of FXTAS, including the data presented at the First International Conference on *FMR1* Premutation (23 to 26 June, Perugia, Italy). For other aspects of FXTAS, there are a number of excellent reviews available [28-31], as well the additional articles published in this special issue of *JND*.

CGG repeats are unstable, and tend to expand over time or with successive generations

Increased CGG repeat numbers are associated with an increased risk of FXTAS and with an increased severity and reduced age of onset of FXTAS symptoms [5,6]. The CGG repeat tract responsible for FXTAS is polymorphic in the human population. Normal alleles have between 6 and 45 repeats, and are relatively stable on intergenerational transmission. However, as the repeat number increases, so too does the likelihood that the repeat tract will expand or gain additional repeat units on intergenerational transfer. AGG interruptions, which are commonly seen within *FMR1* alleles, typically at 10 to 11 and 20 to 21 triplets from the 5′ end [32], are associated with a reduced risk of expansion [33,34]. Contractions do occur [35], but much less frequently. One of the consequences of the expansion bias is that alleles tend to increase in repeat number with successive generations. In addition to intergenerational expansion, somatic expansion is also seen in certain organs, including the brain in mice, and both lymphocytes and brain tissue in humans [27,36]. Somatic expansion may contribute to the repeat-length mosaicism that is seen in some human premutation carriers [27,37-40]. This somatic expansion has the potential to exacerbate FXTAS symptoms, particularly in carriers of alleles with more than 100 CGG repeats, where the repeat may be particularly prone to expansion.

The mechanism responsible for these expansions is unknown. A number of other diseases are known to result from expansions of tracts containing these repeats or other short repeat units. Whether or not these diseases, which are referred to collectively as the repeat expansion diseases [41], share a common expansion mechanism is unknown. However, the unusual nature of these mutations suggests that they might. The expansion bias clearly differentiates the instability in these diseases from the classic microsatellite instability seen in certain cancers, where the repeat is as likely to lose repeats as it is to gain them.

The individual strands of expanded CGG repeats, like other repeats that cause repeat expansion diseases, form secondary structures, including hairpins and quadruplexes [42]. These structures affect DNA processing enzymes such as DNA polymerase, both in vitro [42] and in vivo ([43,44]. It is generally thought that these structures are the trigger or substrate for expansion [41]. However, expansion in brain and liver, which are organs with a low proliferative capacity, along with expansion in mouse oocytes [45,46], suggest that the expansion mechanism may not involve aberrant DNA replication. Rather, given that oxidative stress exacerbates expansion in mice [47], it may be that expansion results from the aberrant repair of oxidized DNA or of DNA that is damaged in other ways. In contrast to generalized microsatellite instability, expansion in premutation mice actually requires *Msh2* [48]. However, whether expansion involves disruption of classic mismatch repair or involves another MSH2-dependent process is unknown. Although *Msh2* is required for expansion, it is not required for contractions [45,46,48]. Thus, it seems likely that expansions and contractions occur by different mechanisms, and the expansion bias seen in FX pedigrees may reflect a more efficient operation of the expansion process relative to the process that generates contractions.

Expression of FMR1 mRNA is increased in premutation carriers

Several studies have shown that premutation alleles are characterized by high mRNA expression levels [13,15,16,49]. In our recent analysis, *FMR1* gene expression levels were measured in peripheral blood leukocytes from a total of 806 males across the whole range of CGG repeats, including normal individuals (n = 463), and individuals carrying intermediate (n = 60) and premutation (n = 283) alleles [27]. The results showed that *FMR1* mRNA levels increased with increased CGG repeat number, and that a significant increase ($P < 0.001$) was detectable for allele lengths as short as 35 CGG repeats (Figure 1). In addition, nuclear run-on experiments indicated that this elevated level of *FMR1* mRNA in premutation carriers is caused by increased transcription efficiency rather than increased mRNA stability [14,15]. Despite higher levels of *FMR1* transcripts, mild deficits of FMRP have been found in premutation carriers, and are probably due to a deficit in translational efficiency, particularly in the upper premutation range [13,16,49]. Thus, an FMRP deficiency is probably not the principal cause of FXTAS. Instead, the crucial observation that RNAs containing expanded CGG repeats accumulate in nuclear RNA aggregates in brain sections of patients with FXTAS [50] supports the notion that elevated levels of *FMR1* mRNA trigger neuronal toxicity. In support of this hypothesis, heterologous expression of 90 CGG repeats in *Drosophila melanogaster* was shown to cause neurodegeneration and formation of ubiquitin inclusions [18]. Similarly, a knock-in (KI) mouse model, in which the endogenous eight CGG repeats of the murine *Fmr1* were

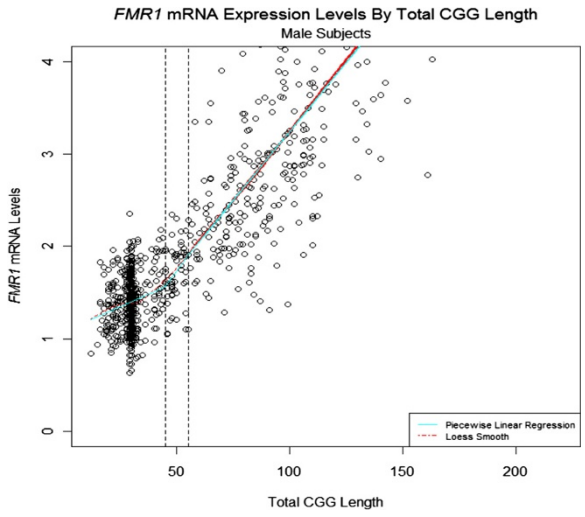


Figure 1 Quantification of *FMR1* mRNA levels in the three allele categories (normal, intermediate, and premutation) shows that *FMR1* mRNA expression increases significantly with increasing CGG repeat number. The solid blue line on the plot shows a piecewise linear regression fit, with *FMR1* mRNA expression increased significantly in all three groups ($P = 0.012$, $P < 0.001$, and $P < 0.001$ for normal, intermediate, and premutation carriers, respectively).

replaced with an expansion containing around 100 CGG repeats of human origin, showed ubiquitin-positive nuclear inclusions, and mild neuromotor and behavioral disturbances [17,51,52]. Finally, expression of transcripts containing 90 CGG repeats in a transgenic mouse model recapitulated some of the neuropathological and molecular features of FXTAS, despite the presence of a normal *Fmr1* allele [19] (see also review on animal models for FXTAS in this issue). These animal models show that the expression of *FMR1* mRNA containing expanded CGG repeats is both necessary and sufficient to cause pathological features characteristic of human FXTAS. Several mechanisms have been proposed to explain how increased expression of a RNA containing expanded CGG repeats could be pathogenic.

Is pathology the result of an RNA gain-of-function mechanism?

The first recognized examples of RNA gain-of-function diseases were two other repeat expansion diseases, myotonic dystrophy type 1 and 2 (DM1 and DM2) [53]. DM is the most common muscular dystrophy in adults, and in this condition, RNAs containing hundreds to thousands of CUG (DM1) or CCUG (DM2) repeats accumulate in nuclear RNA aggregates that sequester the Muscleblind-like (MBNL) splicing factors. Depletion of the free pool of MBNL1 leads to specific alternative splicing changes, which ultimately result in the symptoms of DM [53]. Extending this RNA gain-of-function model to FXTAS, the expanded CGG repeats are predicted to sequester specific proteins, resulting in loss of their normal functions, which would ultimately cause the symptoms of FXTAS [54,55]. Consistent with this idea, Iwahashi and collaborators [56] identified more than 20 proteins from inclusions purified from brains of patients with FXTAS. Of these, two RNA binding proteins were of special interest. The first, hnRNP A2/B1 is mutated in families with inherited degeneration affecting muscle, brain, bone, and motor neurons [57], while the second, MBNL1, is the splicing factor that is involved in DM [58]. However, a role for MBNL1 in FXTAS has been excluded, because no genetic interaction between MBNL1 and CGG-mediated neurodegeneration was observed in the fly model of FXTAS [59], and no misregulation of splicing events regulated by MBNL1 was observed in brain samples from patients with FXTAS [60]. By contrast, binding of hnRNP A2/B1 to RNA containing expanded CGG repeats was confirmed by independent proteomic and in vitro analyses [60,61]. Furthermore, overexpression of hnRNP A2/B1 rescued the neurodegeneration in transgenic *Drosophila* expressing 90 CGG repeats [59,61]. Interaction of hnRNP A2/B1 with RNA containing expanded CGG repeats was evident in cytoplasmic cerebellar lysates. By contrast, nuclear hnRNP A2/B1 presented little binding to CGG RNA, suggesting that some modifications of hnRNP A2/B1, either in the nucleus or in the cytoplasm, may alter the ability of hnRNP A2/B1 to bind to CGG RNA repeats [59]. The importance of titration of the cytoplasmic pool of hnRNP A2/B1 was further demonstrated by expression of expanded CGG repeats in primary cultures of rat sympathetic neurons [62]. RNA containing CGG repeats competed for binding of hnRNP A2/B1 to BC1 RNA, a dendritic regulatory RNA, resulting in impaired dendritic delivery of the BC1 RNA [62]. However, no misregulation of splicing events regulated by nuclear hnRNP A2/B1 was observed in brain samples of patients with FXTAS [60]. Overall, these data suggest that expanded CGG repeats recruit hnRNP A2/B1, resulting in depletion of the cytoplasmic but not the nuclear pool of hnRNP A2/B1. In addition, the ability of hnRNP A proteins to unfold tetraplex RNA structures, formed by expanded CGG repeats [63,64], raises the possibility that hnRNP A2/B1 may also act as a RNA chaperone that destabilizes these RNA structures. Finally, Sofola and collaborators [59] demonstrated that hnRNP A2/B1 recruits, *in trans* and through protein-protein interactions, other proteins such as CUGBP1, an RNA binding protein, whose expression is increased in heart samples of patients with DM [65]. These data indicated that proteins binding to CGG RNA may recruit other proteins, resulting in dynamic aggregates that expand over time, a model later confirmed in COS7 cells expressing 60 CGG repeats [60]. Overexpression of either hnRNP A2/B1 or CUGBP1 rescued neurodegeneration in a *Drosophila* model of FXTAS,

highlighting the potential importance of hnRNP A2/B1 and CUGBP1 to FXTAS pathology [59].

In addition to hnRNP A2/B1, proteomic analyses performed by Jin and collaborators [61] also showed that purine-rich binding protein α (Purα) binds robustly to RNA containing expanded CGG repeats. Purα is a single-stranded cytoplasmic DNA and RNA binding protein that has been implicated in many biological processes, including RNA transport and translation. Importantly, overexpression of Purα rescued neurodegeneration in a *Drosophila* model of FXTAS [61]. However, presence of Purα within nuclear aggregates in FXTAS brain samples is inconsistently observed. Jin et al. found Purα in cytoplasmic inclusions in *Drosophila* expressing 90 CGG repeats, and in inclusions in superior-mid temporal cortex neurons from human FXTAS brain sections [61]. By contrast, Iwashashi et al. did not detect Purα in purified inclusions from cerebral cortex of patients with FXTAS [56]. Furthermore, Purα-positive inclusions have not been observed in mouse models of FXTAS [66], or in hippocampal and cortical brain section of patients with FXTAS [67]. These results suggest that the composition of the inclusions varies from one brain region to the next and from one model organism to the other. Analogous to the recruitment of CUGBP1 by hnRNP A2/B1 to RNA containing expanded CGG repeats, Purα was shown to recruit Rm62, the *Drosophila* ortholog of the RNA helicase P68/DDX5 [68]. Expression of expanded CGG repeats resulted in the post-transcriptional downregulation of Rm62, ultimately resulting in nuclear accumulation of *Hsp70* mRNA and of other mRNAs involved in stress and immune responses [68]. Overexpression of Rm62 rescued neurodegeneration in flies expressing 90 CGG repeats, highlighting the potential importance of P68/DDX5 to FXTAS pathology [68].

SAM68, a splicing regulator encoded by the *KHDRBS1* gene, was also found in CGG RNA aggregates [60]. However, overexpression of SAM68 was not sufficient to rescue neuronal cell death induced by expression of expanded CGG repeats [67]. As with CUGBP1 and Rm62, SAM68 did not bind directly to CGG repeats, and recruitment of SAM68 within CGG RNA aggregates occurred in *trans* through protein-protein interactions [59]. Sellier and collaborators [67] also showed that DROSHA-DGCR8, the enzymatic complex that processes pri-microRNAs into pre-miRNAs, associated specifically with CGG repeats of pathogenic size. Sequestration of DROSHA-DGCR8 within CGG RNA aggregates resulted in reduced processing of pri-miRNAs in cells expressing expanded CGG repeats, and in brain samples from patients with FXTAS. Overexpression of DGCR8 rescued neuronal cell death induced by expression of expanded CGG repeats [67]. These results suggest that titration of DGCR8 by expanded CGG repeats is a leading event to CGG-induced neuronal cell death. However, recent analyses of miRNA expression in blood samples of patients with FXTAS and in *Drosophila* expressing CGG repeats did not show a global downregulation of miRNA, but rather, the expression of some specific miRNAs was misregulated [69,70]. Whether depletion of DROSHA-DGCR8 varies in blood and brain of patients with FXTAS, and whether the Drosha-Pasha complex is sequestered in cytoplasmic aggregates in *Drosophila* expressing expanded CGG repeats, remains to be determined. Similarly, whether overexpression of hnRNP A2/B1, P68/DDX5, DROSHA-DGCR8, or CUGBP1 rescues any phenotype in mouse models expressing expanded CGG repeats, would be necessary to determine the importance of these candidate proteins to FXTAS pathology.

These caveats aside, the observations described above suggest that CGG repeats could be pathogenic by sequestering specific RNA binding proteins, resulting in loss of their normal functions, and thus lead to neuronal cell dysfunction (Figure 2) [56,58,61,67,68]. However, this attractive model has some weaknesses. First, the inclusions observed in FXTAS brain sections differ from those seen in DM, a typical RNA gain-of-function disorder. In FXTAS, inclusions are larger and ubiquitinated, and contain various chaperone proteins such as Hsp27, Hsp70, and αB-crystallin [9,56]. In short, these large inclusions resemble the aggregates seen in protein-mediated disorders, although they are negative for the typical proteins found in tauopathies, synucleinopathies, or polyQ disorders (for example, Huntington's disease). Second, and most disconcerting, although inclusions in brain samples of patients with FXTAS contain the mutant *FMR1* RNA with expanded CGG repeats [50], a mouse model, in which the endogenous eight CGG repeats of *Fmr1* is replaced with an expansion containing around 100 CGG repeats, shows numerous ubiquitin inclusions but only rare aggregates of SAM68 or DROSHA-DGCR8, associated with rare RNA aggregates of expanded CGG repeats [60,67]. Similarly, overexpression of expanded CGG repeats leads to formation of nuclear RNA aggregate in some cell types, including primary cultures of hippocampal embryonic mouse neurons and PC12, COS7, and SKOV3 immortalized cell lines, but no RNA aggregates have been observed in A172, U-937, THP1, HeLa, HEK293, NG108-15, IMR-32, Neuro-2a, SH-SY5Y, SK-N-MC, or SK-N-SH cells [22,60]. In short, not all cell lines can support CGG repeat aggregate formation, whereas in DM, expression of expanded CUG or CCUG repeats consistently results in formation of RNA foci. Thirdly, a recent and provocative study demonstrates that the toxic effect of CGG repeats depends on their location [71]. Moving expanded CGG repeats from a 5′ UTR to a 3′ UTR position reduced their toxic effect in *Drosophila*, whereas expanded CUG or CCUG repeats were found

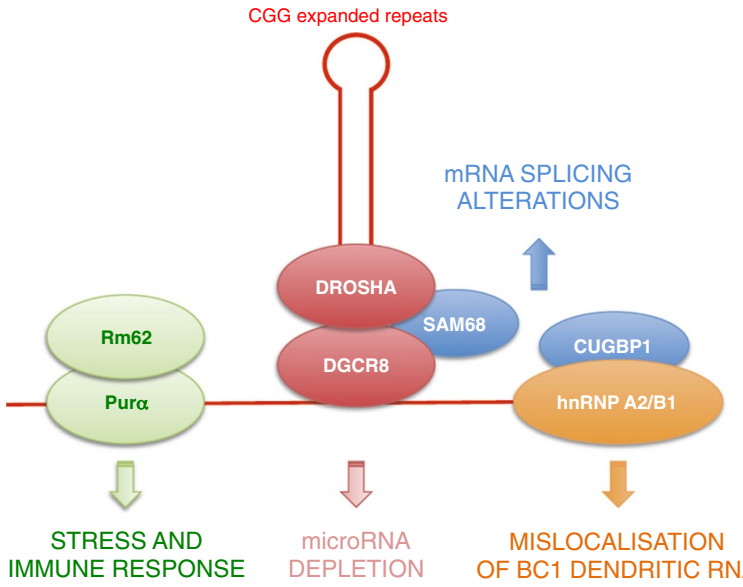


Figure 2 Various RNA binding proteins have been found to associate with RNAs containing expanded CGG repeats. Purα, DGCR8, and hnRNP A2/B1 bind directly to the CGG-containing RNA, whereas Rm62, SAM68, and CUGBP1 are recruited in *trans* through protein-protein interactions.

to be pathogenic in whatever location tested, provided they were expressed in sufficient amounts to deplete MBNL proteins. These data led Todd and collaborators to reconsider the model of RNA binding protein sequestration, and to explore further the molecular mechanisms of FXTAS.

Non-canonical AUG translation produces a polyglycine-containing protein in FXTAS

An unexpected observation, made by Todd and colleagues in flies transgenic for a construct containing 90 expanded CGG repeats cloned upstream of the green fluorescent protein (GFP) cDNA, was that some of the GFP signal was found in cytoplasmic inclusions. Western blotting analysis showed a band of the expected size for GFP, but also detected a protein 12 kDa larger [71]. Because translation of expanded CAG repeats in the absence of an ATG initiation codon (repeat-associated non-ATG translation or RAN translation) had been previously reported [72], Todd and collaborators tested whether expanded CGG repeats could be translated despite the fact that no ATG codon is present upstream of the repeats. Their analysis revealed that, indeed, translation of CGG repeats occurs in two out of the three frames, giving rise to short proteins containing either a polyalanine or a polyglycine stretch. Expression of the polyglycine protein resulted in the formation of protein inclusions, which were toxic both in neuronal transfected cells and in *Drosophila*. Further analyses of the polyglycine protein revealed that its translation was probably initiated at non-canonical AUG codons, such as CUG and GUG, which were located upstream of the CGG repeats. A role for non-canonical translation initiation in inclusion formation is consistent with data from two different KI mice mouse models. In a mouse model that showed numerous ubiquitin inclusions, the expanded CGG repeat from a human premutation allele was cloned, along with sequences upstream of the CGG repeats in humans that contained the non-AUG initiations codons [17]. By contrast, in a second mouse model, in which the mouse 5′ flanking sequence was retained, a stop codon was found to be located just upstream of the expanded CGG repeats [21]. These latter mice showed relatively few ubiquitinated aggregates, thus supporting the notion that non-ATG-initiated translation of the CGG tract is required to generate most of the inclusions [71]. That this unusual mode of translation may play a role in FXTAS is evidenced by the fact that, with the aid of specific antibodies, polyglycine protein can be seen in brain sections of patients with FXTAS [71]. Overall, these observations suggest that a protein gain of function may also occur in cells of patients with FXTAS. However, what contribution the polyglycine-containing or polyalanine-containing proteins make to the etiology of FXTAS is an exciting open question.

Non-coding transcription of the FMR1 locus: a role in FMR1 mRNA toxicity?

The majority of the human genome is transcribed but not translated. Such RNAs are classified as long non-coding RNAs (lncRNAs) when longer than 200 nucleotides [73-75]. To date, relatively few lncRNAs have been functionally characterized, but increasing evidence suggests that many may have important functions, including

the regulation of transcription, RNA processing and translation, DNA methylation, and chromatin architecture, both locally (*cis*-acting) and across some genomic distance (*trans*-acting) [76-78].

In addition to the *FMR1* transcript, a variety of RNAs are produced from the *FMR1* locus. Therefore, it is possible that these lncRNAs produced from the *FMR1* locus may modulate certain aspects of FXS/FXTAS, as has been shown in other human diseases [79]. For example, Kumari and Usdin described an abundant antisense transcript of about 5 kb that spans the region upstream of the FMR1 promoter, and whose expression does not change in response to repeat expansion [80]. By contrast Ladd and coworkers described a transcript, Antisense *FMR1 (ASFMR1*), that spans the expanded CGG repeats and whose expression is elevated in lymphoblastoid cells and peripheral blood leukocytes of individuals with premutation alleles, while it is not expressed in those with full mutation alleles [81]. Multiple splice forms of *ASFMR1* have been identified, which show differential expression in carriers of premutation and normal alleles [81]. One of these *ASFMR1* splice variants contains a small intron that uses a non-consensus CT-AC splice site that is transcribed in a premutation cell line, but is absent in a normal cell line [81]. We compared the expression levels of this *ASFMR1* isoform in blood from individuals with alleles ranging from normal to premutation, and found a significant increase with CGG repeat number ($P < 0.001$) (Figure 3) [27]. Of interest, both unspliced and spliced *ASFMR1* transcripts contain putative open-reading frames encoding polyproline peptides, resulting from antisense-oriented translation of the expanded CGG repeats [81]. Whether *ASFMR1* containing expanded CCG repeats is translated and participates in the formation of the pathogenic nuclear inclusions observed in patients with FXTAS remain to be tested.

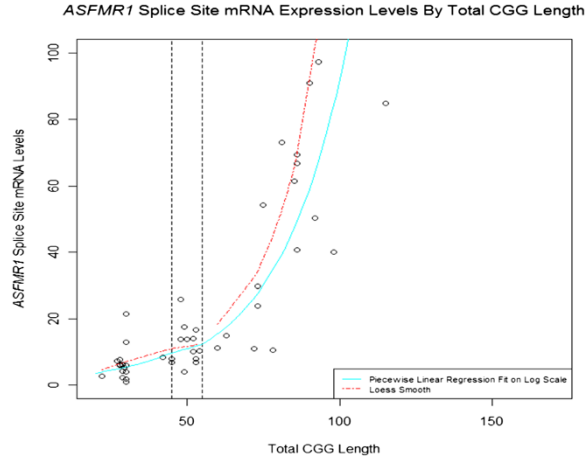


Figure 3 Expression of the minor splice isoform in the *ASFMR1* transcript (131 bp), located near the *ASFMR1* promoter. Expression of this isoform increases in premutation carriers ($P < 0.001$), and shows a similar trend in subjects with intermediate alleles ($P = 0.0528$) compared with normal alleles. The solid blue line on the plot shows a piecewise linear regression fit (fitted on the log scale then exponentiated for plotting), with separate slopes in the normal, intermediate, and premutation alleles.

Another antisense transcript, *FMR4*, originates upstream of the *FMR1* start site, and covers 2.4 kb of sequence [82]. *FMR4* is widely expressed in fetal and adult human tissues, and throughout human and macaque brain regions. Expression of *FMR4*, like that of *ASFMR1* and *FMR1*, is increased in brain tissue of premutation individuals and is silenced in individuals with the full mutation [82]. Importantly, *FMR4* overexpression was shown to increase cell proliferation, whereas *FMR4* downregulation induced apoptosis in vitro [82]. Additionally, no *cis*-acting effect was observed upon expression of the *FMR1* gene. Therefore, it was hypothesized that *FMR4* influences proliferation pathways in *trans*, by targeting distal genomic loci. Current work is focused on defining a role for this transcript, as it has been found to affect the chromatin state and transcription of several genes involved in neuronal differentiation, axon guidance, and synaptic signaling, as well as cell cycle regulators (Peschansky and Pastori, unpublished data).

Two new transcripts arising from the *FMR1* locus, *FMR5* and *FMR6,* were recently identified [83]. *FMR5* is a sense-oriented lncRNA transcribed from approximately 1 kb upstream of the *FMR1* transcription start site (TSS). *FMR5* is not differentially expressed in human brain from unaffected individuals compared with full mutation and premutation patients, suggesting that its transcription is independent of CGG repeat expansion. Furthermore, the TSS of *FMR5* appears not to be affected by the chromatin silencing that occurs within full mutation alleles, or by the open chromatin hypothesized to increase transcription of *FMR1* premutation alleles. *FMR6* is a spliced long antisense transcript, 600 nucleotides i, whose sequence is entirely complementary to the 3′ region of *FMR1* [83]. It begins in the 3′UTR, ends in exon 15 of *FMR1*, and uses the same splice junctions as *FMR1*. An unexpected finding was that *FMR6* is reduced in premutation carriers, suggesting that abnormal transcription and/or chromatin remodeling occurs toward the distal end of the locus. However, the chromatin marks associated with the 3′ end of *FMR1* in premutation carriers have yet to be described. The function of *FMR6* remains to be identified, but its complementarity to the 3′ region of *FMR1* presents several interesting possibilities. *FMR6* may bind to *FMR1* mRNA, thereby regulating the stability, splicing, subcellular localization, or translational efficiency of *FMR1*, as has been described for other lncRNAs [77]. Notably, *FMR6* overlaps miR-19a and miR-19b binding sites in the *FMR1* 3′ UTR [84], suggesting that *FMR6* may modulate the stability or translational efficiency of *FMR1* by interfering with

microRNA binding. Overall, these results highlight the importance of non-coding transcription of the *FMR1* locus. However, much work remains to fully understand the relevance of these transcripts to the pathology observed in premutation carriers.

Conclusion

The restriction of FXTAS clinical features to unmethylated, transcriptionally active alleles with large CGG repeat numbers suggests that the expression of a mutant RNA is pathogenic to neuronal cells [55]. This hypothesis is supported by data from cell, fly, and mouse models [17-22]. However, how these RNAs cause neuronal cell dysfunction and FXTAS symptoms remains unclear. One model proposes that the RNA containing expanded CGG repeats is pathogenic via its sequestration of specific RNA binding proteins. Various proteins, including Purα, Rm62, CUGBP1, hnRNP A2/B1, SAM68, and DROSHA-DGCR8, have been shown to bind directly or through a protein partner to expanded CGG repeats [56,59,61,67,68]. However, it remains to be tested whether overexpression of these candidate proteins rescues any phenotype in mouse models expressing expanded CGG repeats. A second mechanism involves non-canonical translation initiation of expanded CGG repeats. resulting in expression of toxic polyglycine-containing and polyalanine-containing proteins [71]; however, how these proteins promote neuronal cell dysfunction is an open question. A third model is associated with the expression of antisense *FMR1* transcripts. Further investigation is needed to evaluate the pathological consequences of expression of *ASFMR1* or other long non-coding RNA mapping within the *FMR1* gene. Finally, although decreased expression of FMRP is probably not the principal cause of FXTAS, it cannot be excluded that a reduction in FMRP plays a role in modulating some of FXTAS features. In that context, the level of FMRP depletion in brain samples from a larger cohort of patients with FXTAS needs to be measured.

In conclusion, in addition to increased *FMR1* mRNA production, protein titration, non-AUG translation,

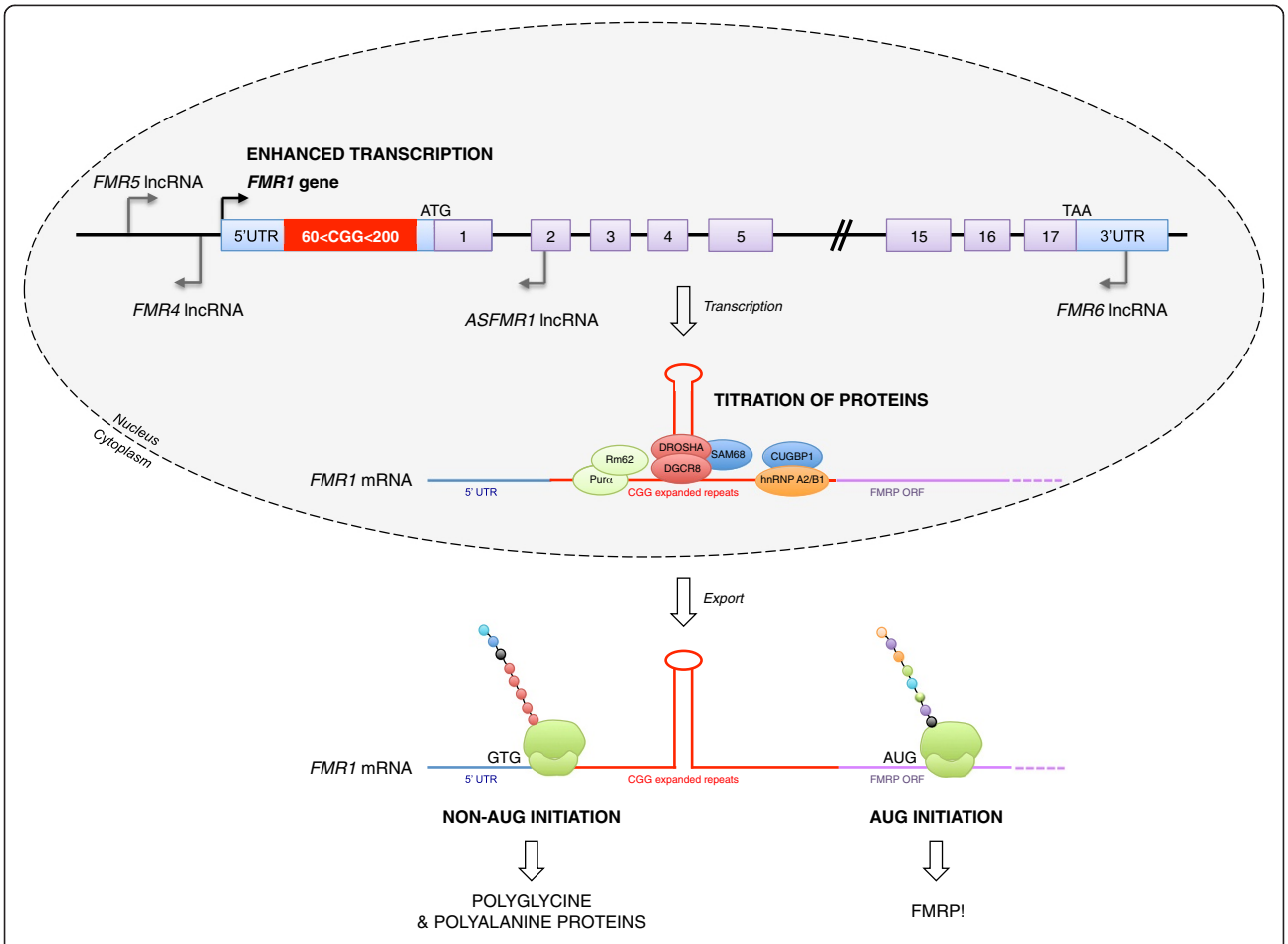


Figure 4 Mechanisms that may contribute to fragile X-associated tremor/ataxia (FXTAS) pathology. Expression of *FMR1* mRNA and associated antisense transcripts are increased inpremutation carriers. Some transcripts accumulate in the nucleus and recruit various RNA binding proteins, while export and non-AUG translation results in production of polyglycine-containing or polyalanine-containing proteins.

antisense transcription, and decreased expression of FMRP are a number of non-exclusive mechanisms that may all contribute to FXTAS pathology. It is possible that contributions to pathology from more than one mechanism may help to explain the great variability in clinical presentation of premutation individuals, aspects of which have heretofore not been accounted for by CGG expansion size, mosaicism, methylation, alternative spliced isoforms, additional genomic changes, or other known factors. Thus, more work is needed to determine the relative contribution of these processes to disease pathology in this multifaceted disorder (Figure 4).

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
CS, KU, CP, VJP, FT and NCB wrote the paper. All authors read and approved the final manuscript.

Acknowledgments
We thank Paul Hagerman and Peter Todd for invaluable and fruitful discussions. We sincerely apologize to all colleagues whose work could not be cited due to space limitations. This work was supported by INSERM (NCB), ANR and E-RARE 'CURE FXTAS' (NCB), ERC 'RNA DISEASES' (NCB), the National Institute of Mental Health (grant number 5R01MH084880-05) and the Intramural Program of the National Institute of Diabetes and Digestive and Kidney Diseases (NIH; DK057808-05).

Author details
[1]Department of Translational Medicine, IGBMC, INSERM U964 Illkirch, France. [2]Section on Gene Structure and Disease, NIDDK, National Institutes of Health, Bethesda MD 20892, USA. [3]Department of Psychiatry and Behavioral Sciences and Center for Therapeutic Innovation, Hussman Institute for Human Genomics, University of Miami, Miller School of Medicine, Miami FL 33136, USA. [4]Department of Biochemistry and Molecular Medicine, University of California, Davis, School of Medicine, Sacramento CA 95817, USA. [5]MIND Institute, University of California Davis Medical Center, Sacramento CA 95817, USA. [6]Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS UMR7104, INSERM U964, University of Strasbourg, 1 rue Laurent Fries, Illkirch F-67404, France.

Received: 14 October 2013 Accepted: 15 November 2013
Published: 30 July 2014

References
1. Hagerman RJ, Leehey M, Heinrichs W, Tassone F, Wilson R, Hills J, Grigsby J, Gage B, Hagerman PJ: **Intention tremor, parkinsonism, and generalized brain atrophy in male carriers of fragile X.** *Neurology* 2001, **57:**127–130.
2. Tassone F, Iong KP, Tong TH, Lo J, Gane LW, Berry-Kravis E, Nguyen D, Mu LY, Laffin J, Bailey DB, Hagerman RJ: **FMR1 CGG allele size and prevalence ascertained through newborn screening in the United States.** *Genome Med* 2012, **4:**100.
3. Maenner MJ, Baker MW, Broman KW, Tian J, Barnes JK, Atkins A, McPherson E, Hong J, Brilliant MH, Mailick MR: **FMR1 CGG expansions: prevalence and sex ratios.** *Am J Med Genet B Neuropsychiatr Genet* 2013, **162B:**466–473.
4. Jacquemont S, Hagerman RJ, Leehey MA, Hall DA, Levine RA, Brunberg JA, Zhang L, Jardini T, Gane LW, Harris SW, Herman K, Grigsby J, Greco CM, Berry-Kravis E, Tassone F, Hagerman PJ: **Penetrance of the fragile X-associated tremor/ataxia syndrome in a premutation carrier population.** *Jama* 2004, **291:**460–469.
5. Tassone F, Adams J, Berry-Kravis EM, Cohen SS, Brusco A, Leehey MA, Li L, Hagerman RJ, Hagerman PJ: **CGG repeat length correlates with age of onset of motor signs of the fragile X-associated tremor/ataxia syndrome (FXTAS).** *Am J Med Genet B Neuropsychiatr Genet* 2007, **144B:**566–569.
6. Leehey MA, Berry-Kravis E, Goetz CG, Zhang L, Hall DA, Li L, Rice CD, Lara R, Cogswell J, Reynolds A, Gane L, Jacquemont S, Tassone F, Grigsby J, Hagerman RJ, Hagerman PJ: **FMR1 CGG repeat length predicts motor dysfunction in premutation carriers.** *Neurology* 2008, **70:**1397–1402.
7. Jacquemont S, Hagerman RJ, Leehey M, Grigsby J, Zhang L, Brunberg JA, Greco C, Des Portes V, Jardini T, Levine R, Berry-Kravis E, Brown WT, Schaeffer S, Kissel J, Tassone F, Hagerman PJ: **Fragile X premutation tremor/ataxia syndrome: molecular, clinical, and neuroimaging correlates.** *Am J Hum Genet* 2003, **72:**869–878.
8. Brunberg JA, Jacquemont S, Hagerman RJ, Berry-Kravis EM, Grigsby J, Leehey MA, Tassone F, Brown WT, Greco CM, Hagerman PJ: **Fragile X premutation carriers: characteristic MR imaging findings of adult male patients with progressive cerebellar and cognitive dysfunction.** *AJNR Am J Neuroradiol* 2002, **23:**1757–1766.
9. Greco CM, Hagerman RJ, Tassone F, Chudley AE, Del Bigio MR, Jacquemont S, Leehey M, Hagerman PJ: **Neuronal intranuclear inclusions in a new cerebellar tremor/ataxia syndrome among fragile X carriers.** *Brain* 2002, **125:**1760–1771.
10. Greco CM, Berman RF, Martin RM, Tassone F, Schwartz PH, Chang A, Trapp BD, Iwahashi C, Brunberg J, Grigsby J, Hessl D, Becker EJ, Papazian J, Leehey MA, Hagerman RJ, Hagerman PJ: **Neuropathology of fragile X-associated tremor/ataxia syndrome (FXTAS).** *Brain* 2006, **129:**243–255.
11. Greco CM, Soontrapornchai K, Wirojanan J, Gould JE, Hagerman PJ, Hagerman RJ: **Testicular and pituitary inclusion formation in fragile X associated tremor/ataxia syndrome.** *J Urol* 2007, **177:**1434–1437.
12. Hunsaker MR, Greco CM, Spath MA, Smits AP, Navarro CS, Tassone F, Kros JM, Severijnen LA, Berry-Kravis EM, Berman RF, Hagerman PJ, Willemsen R, Hagerman RJ, Hukema RK: **Widespread non-central nervous system organ pathology in fragile X premutation carriers with fragile X-associated tremor/ataxia syndrome and CGG knock-in mice.** *Acta Neuropathol* 2011, **122:**467–479.
13. Kenneson A, Zhang F, Hagedorn CH, Warren ST: **Reduced FMRP and increased FMR1 transcription is proportionally associated with CGG repeat number in intermediate-length and premutation carriers.** *Hum Mol Genet* 2001, **10:**1449–1454.
14. Tassone F, Beilina A, Carosi C, Albertosi S, Bagni C, Li L, Glover K, Bentley D, Hagerman PJ: **Elevated FMR1 mRNA in premutation carriers is due to increased transcription.** *RNA* 2007, **13:**555–562.
15. Tassone F, Hagerman RJ, Taylor AK, Gane LW, Godfrey TE, Hagerman PJ: **Elevated levels of FMR1 mRNA in carrier males: a new mechanism of involvement in the fragile-X syndrome.** *Am J Hum Genet* 2000, **66:**6–15.
16. Peprah E, He W, Allen E, Oliver T, Boyne A, Sherman SL: **Examination of FMR1 transcript and protein levels among 74 premutation carriers.** *J Hum Genet* 2010, **55:**66–68.
17. Willemsen R, Hoogeveen-Westerveld M, Reis S, Holstege J, Severijnen LA, Nieuwenhuizen IM, Schrier M, van Unen L, Tassone F, Hoogeveen AT, Hagerman PJ, Mientjes EJ, Oostra BA: **The FMR1 CGG repeat mouse displays ubiquitin-positive intranuclear neuronal inclusions; implications for the cerebellar tremor/ataxia syndrome.** *Hum Mol Genet* 2003, **12:**949–959.
18. Jin P, Zarnescu DC, Zhang F, Pearson CE, Lucchesi JC, Moses K, Warren ST: **RNA-mediated neurodegeneration caused by the fragile X premutation rCGG repeats in Drosophila.** *Neuron* 2003, **39:**739–747.
19. Hashem V, Galloway JN, Mori M, Willemsen R, Oostra BA, Paylor R, Nelson DL: **Ectopic expression of CGG containing mRNA is neurotoxic in mammals.** *Hum Mol Genet* 2009, **18:**2443–2451.
20. Handa V, Goldwater D, Stiles D, Cam M, Poy G, Kumari D, Usdin K: **Long CGG-repeat tracts are toxic to human cells: implications for carriers of Fragile X premutation alleles.** *FEBS Lett* 2005, **579:**2702–2708.
21. Entezam A, Biacsi R, Orrison B, Saha T, Hoffman GE, Grabczyk E, Nussbaum RL, Usdin K: **Regional FMRP deficits and large repeat expansions into the full mutation range in a new Fragile X premutation mouse model.** *Gene* 2007, **395:**125–134.
22. Arocena DG, Iwahashi CK, Won N, Beilina A, Ludwig AL, Tassone F, Schwartz PH, Hagerman PJ: **Induction of inclusion formation and disruption of lamin A/C structure by premutation CGG-repeat RNA in human cultured neural cells.** *Hum Mol Genet* 2005, **14:**3661–3671.
23. Hall D, Tassone F, Klepitskaya O, Leehey M: **Fragile X-associated tremor ataxia syndrome in FMR1 gray zone allele carriers.** *Mov Disord* 2012, **27:**296–300.
24. Liu Y, Winarni TI, Zhang L, Tassone F, Hagerman RJ: **Fragile X-associated tremor/ataxia syndrome (FXTAS) in grey zone carriers.** *Clin Genet* 2013, **84:**74–77.

25. Loesch DZ, Sherwell S, Kinsella G, Tassone F, Taylor A, Amor D, Sung S, Evans A: **Fragile X-associated tremor/ataxia phenotype in a male carrier of unmethylated full mutation in the FMR1 gene.** *Clin Genet* 2012, **82:**88–92.
26. Santa María L, Pugin A, Alliende M, Aliaga S, Curotto B, Aravena T, Tang HT, Mendoza-Morales G, Hagerman R, Tassone F: **Fxtas in an unmethylated mosaic male with fragile X syndrome from chile.** *Clin Genet* 2013. doi: 10.1111/cge.12278. [Epub ahead of print].
27. Pretto DI, Tang HT, Lo J, Morales G, Hagerman R, Hagerman PJ, Tassone F: *CGG-repeat number and methylation instability in premutation alleles.* Perugia, Italy: First International Conference on FMR1 Premutation 23 to 26 June 2013; 2013.
28. Willemsen R, Levenga J, Oostra BA: **CGG repeat in the FMR1 gene: size matters.** *Clin Genet* 2011, **80:**214–225.
29. Li Y, Jin P: **RNA-mediated neurodegeneration in fragile X-associated tremor/ataxia syndrome.** *Brain Res* 2012, **1462:**112–117.
30. Hagerman R, Hagerman P: **Advances in clinical and molecular understanding of the FMR1 premutation and fragile X-associated tremor/ataxia syndrome.** *Lancet Neurol* 2013, **12:**786–798.
31. Garcia-Arocena D, Hagerman PJ: **Advances in understanding the molecular basis of FXTAS.** *Hum Mol Genet* 2010, **19:**R83–R89.
32. Eichler EE, Holden JJ, Popovich BW, Reiss AL, Snow K, Thibodeau SN, Richards CS, Ward PA, Nelson DL: **Length of uninterrupted CGG repeats determines instability in the FMR1 gene.** *Nat Genet* 1994, **8:**88–94.
33. Nolin SL, Sah S, Glicksman A, Sherman SL, Allen E, Berry-Kravis E, Tassone F, Yrigollen C, Cronister A, Jodah M, Ersalesi N, Dobkin C, Brown WT, Shroff R, Latham GJ, Hadd AG: **Fragile X AGG analysis provides new risk predictions for 45–69 repeat alleles.** *Am J Med Genet A* 2013, **161:**771–778.
34. Yrigollen CM, Durbin-Johnson B, Gane L, Nelson DL, Hagerman R, Hagerman PJ, Tassone F: **AGG interruptions within the maternal FMR1 gene reduce the risk of offspring with fragile X syndrome.** *Genet Med* 2012, **14:**729–736.
35. Tabolacci E, Pomponi MG, Pietrobono R, Chiurazzi P, Neri G: **A unique case of reversion to normal size of a maternal premutation FMR1 allele in a normal boy.** *Eur J Hum Genet* 2008, **16:**209–214.
36. Lokanga RA, Entezam A, Kumari D, Yudkin D, Qin M, Smith CB, Usdin K: **Somatic expansion in mouse and human carriers of fragile X premutation alleles.** *Hum Mutat* 2013, **34:**157–166.
37. Grasso M, Faravelli F, Lo Nigro C, Chiurazzi P, Sperandeo MP, Argusti A, Pomponi MG, Lecora M, Sebastio GF, Perroni L, Andria G, Neri G, Bricarelli FD: **Mosaicism for the full mutation and a microdeletion involving the CGG repeat and flanking sequences in the FMR1 gene in eight fragile X patients.** *Am J Med Genet* 1999, **85:**311–316.
38. Govaerts LC, Smit AE, Saris JJ, VanderWerf F, Willemsen R, Bakker CE, De Zeeuw CI, Oostra BA: **Exceptional good cognitive and phenotypic profile in a male carrying a mosaic mutation in the FMR1 gene.** *Clin Genet* 2007, **72:**138–144.
39. Hantash FM, Goos DG, Tsao D, Quan F, Buller-Burckle A, Peng M, Jarvis M, Sun W, Strom CM: **Qualitative assessment of FMR1 (CGG)n triplet repeat status in normal, intermediate, premutation, full mutation, and mosaic carriers in both sexes: implications for fragile X syndrome carrier and newborn screening.** *Genet Med* 2010, **12:**162–173.
40. Nolin SL, Glicksman A, Houck GE Jr, Brown WT, Dobkin CS: **Mosaicism in fragile X affected males.** *Am J Med Genet* 1994, **51:**509–512.
41. Mirkin SM: **Expandable DNA repeats and human disease.** *Nature* 2007, **447:**932–940.
42. Usdin K, Woodford KJ: **CGG repeats associated with DNA instability and chromosome fragility form structures that block DNA synthesis in vitro.** *Nucleic Acids Res* 1995, **23:**4202–4209.
43. Voineagu I, Surka CF, Shishkin AA, Krasilnikova MM, Mirkin SM: **Replisome stalling and stabilization at CGG repeats, which are responsible for chromosomal fragility.** *Nat Struct Mol Biol* 2009, **16:**226–228.
44. Gerhardt J, Tomishima MJ, Zaninovic N, Colak D, Yan Z, Zhan Q, Rosenwaks Z, Jaffrey SR, Schildkraut CL: *The DNA replication program is altered at the FMR1 locus in fragile X embryonic stem cells.* Perugia, Italy: First International Conference on FMR1 Premutation 23 to 26 June 2013; 2013.
45. Entezam A, Usdin K: **ATR protects the genome against CGG. CCG-repeat expansion in Fragile X premutation mice.** *Nucleic Acids Res* 2008, **36:**1050–1056.
46. Entezam A, Usdin K: **ATM and ATR protect the genome against two different types of tandem repeat instability in Fragile X premutation mice.** *Nucleic Acids Res* 2009, **37:**6371–6377.
47. Entezam A, Lokanga AR, Le W, Hoffman G, Usdin K: **Potassium bromate, a potent DNA oxidizing agent, exacerbates germline repeat expansion in a fragile X premutation mouse model.** *Hum Mutat* 2010, **31:**611–616.
48. Lokanga RA, Zhao X-N, Usdin K: **The mismatch repair protein MSH2 is rate-limiting for repeat expansion in a Fragile X premutation mouse model.** *Hum Mutat* 2013. doi: 10.1002/humu.22464. [Epub ahead of print].
49. Allen EG, Sherman S, Abramowitz A, Leslie M, Novak G, Rusin M, Scott E, Letz R: **Examination of the effect of the polymorphic CGG repeat in the FMR1 gene on cognitive performance.** *Behav Genet* 2005, **35:**435–445.
50. Tassone F, Iwahashi C, Hagerman PJ: **FMR1 RNA within the intranuclear inclusions of fragile X-associated tremor/ataxia syndrome (FXTAS).** *RNA Biol* 2004, **1:**103–105.
51. Brouwer JR, Huizer K, Severijnen LA, Hukema RK, Berman RF, Oostra BA, Willemsen R: **CGG-repeat length and neuropathological and molecular correlates in a mouse model for fragile X-associated tremor/ataxia syndrome.** *J Neurochem* 2008, **107:**1671–1682.
52. Van Dam D, Errijgers V, Kooy RF, Willemsen R, Mientjes E, Oostra BA, De Deyn PP: **Cognitive decline, neuromotor and behavioural disturbances in a mouse model for fragile-X-associated tremor/ataxia syndrome (FXTAS).** *Behav Brain Res* 2005, **162:**233–239.
53. Ranum LP, Cooper TA: **RNA-mediated neuromuscular disorders.** *Annu Rev Neurosci* 2006, **29:**259–277.
54. Nelson DL, Orr HT, Warren ST: **The unstable repeats–three evolving faces of neurological disease.** *Neuron* 2013, **77:**825–843.
55. Hagerman PJ, Hagerman RJ: **The fragile-X premutation: a maturing perspective.** *Am J Hum Genet* 2004, **74:**805–816.
56. Iwahashi CK, Yasui DH, An HJ, Greco CM, Tassone F, Nannen K, Babineau B, Lebrilla CB, Hagerman RJ, Hagerman PJ: **Protein composition of the intranuclear inclusions of FXTAS.** *Brain* 2006, **129:**256–271.
57. Kim HJ, Kim NC, Wang YD, Scarborough EA, Moore J, Diaz Z, MacLea KS, Freibaum B, Li S, Molliex A, Kanagaraj AP, Carter R, Boylan KB, Wojtas AM, Rademakers R, Pinkus JL, Greenberg SA, Trojanowski JQ, Traynor BJ, Smith BN, Topp S, Gkazi AS, Miller J, Shaw CE, Kottlors M, Kirschner J, Pestronk A, Li YR, Ford AF, Gitler AD, *et al*: **Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS.** *Nature* 2013, **495:**467–473.
58. Kanadia RN, Johnstone KA, Mankodi A, Lungu C, Thornton CA, Esson D, Timmers AM, Hauswirth WW, Swanson MS: **A muscleblind knockout model for myotonic dystrophy.** *Science* 2003, **302:**1978–1980.
59. Sofola OA, Jin P, Qin Y, Duan R, Liu H, de Haro M, Nelson DL, Botas J: **RNA-binding proteins hnRNP A2/B1 and CUGBP1 suppress fragile X CGG premutation repeat-induced neurodegeneration in a Drosophila model of FXTAS.** *Neuron* 2007, **55:**565–571.
60. Sellier C, Rau F, Liu Y, Tassone F, Hukema RK, Gattoni R, Schneider A, Richard S, Willemsen R, Elliott DJ, Hagerman PJ, Charlet-Berguerand N: **Sam68 sequestration and partial loss of function are associated with splicing alterations in FXTAS patients.** *Embo J* 2010, **29:**1248–1261.
61. Jin P, Duan R, Qurashi A, Qin Y, Tian D, Rosser TC, Liu H, Feng Y, Warren ST: **Pur alpha binds to rCGG repeats and modulates repeat-mediated neurodegeneration in a Drosophila model of fragile X tremor/ataxia syndrome.** *Neuron* 2007, **55:**556–564.
62. Muslimov IA, Patel MV, Rose A, Tiedge H: **Spatial code recognition in neuronal RNA targeting: role of RNA-hnRNP A2 interactions.** *J Cell Biol* 2011, **194:**441–457.
63. Khateb S, Weisman-Shomer P, Hershco I, Loeb LA, Fry M: **Destabilization of tetraplex structures of the fragile X repeat sequence (CGG)n is mediated by homolog-conserved domains in three members of the hnRNP family.** *Nucleic Acids Res* 2004, **32:**4145–4154.
64. Ofer N, Weisman-Shomer P, Shklover J, Fry M: **The quadruplex r(CGG)n destabilizing cationic porphyrin TMPyP4 cooperates with hnRNPs to increase the translation efficiency of fragile X premutation mRNA.** *Nucleic Acids Res* 2009, **37:**2712–2722.
65. Kuyumcu-Martinez NM, Wang GS, Cooper TA: **Increased steady-state levels of CUGBP1 in myotonic dystrophy 1 are due to PKC-mediated hyperphosphorylation.** *Mol Cell* 2007, **28:**68–78.
66. Galloway JN, Nelson DL: **Evidence for RNA-mediated toxicity in the fragile X-associated tremor/ataxia syndrome.** *Future Neurol* 2009, **4:**785.
67. Sellier C, Freyermuth F, Tabet R, Tran T, He F, Ruffenach F, Alunni V, Moine H, Thibault C, Page A, Tassone F, Willemsen R, Disney MD, Hagerman PJ, Todd PK, Charlet-Berguerand N: **Sequestration of DROSHA and DGCR8 by**

expanded CGG RNA repeats alters microRNA processing in fragile X-associated tremor/ataxia syndrome. *Cell Rep* 2013, **3**:869–880.
68. Qurashi A, Li W, Zhou JY, Peng J, Jin P: **Nuclear accumulation of stress response mRNAs contributes to the neurodegeneration caused by Fragile X premutation rCGG repeats.** *PLoS Genet* 2011, **7**:e1002102.
69. Alvarez-Mora MI, Rodriguez-Revenga L, Madrigal I, Torres-Silva F, Mateu-Huertas E, Lizano E, Friedländer MR, Martí E, Estivill X, Milà M: **MicroRNA expression profiling in blood from fragile X-associated tremor/ataxia syndrome patients.** *Genes Brain Behav* 2013, **12**:595–603.
70. Tan H, Poidevin M, Li H, Chen D, Jin P: **MicroRNA-277 modulates the neurodegeneration caused by Fragile X premutation rCGG repeats.** *PLoS Genet* 2012, **8**:e1002681.
71. Todd PK, Oh SY, Krans A, He F, Sellier C, Frazer M, Renoux AJ, Chen KC, Scaglione KM, Basrur V, Elenitoba-Johnson K, Vonsattel JP, Louis ED, Sutton MA, Taylor JP, Mills RE, Charlet-Berguerand N, Paulson HL: **CGG repeat-associated translation mediates neurodegeneration in fragile X tremor ataxia syndrome.** *Neuron* 2013, **78**:440–455.
72. Zu T, Gibbens B, Doty NS, Gomes-Pereira M, Huguet A, Stone MD, Margolis J, Peterson M, Markowski TW, Ingram MA, Nan Z, Forster C, Low WC, Schoser B, Somia NV, Clark HB, Schmechel S, Bitterman PB, Gourdon G, Swanson MS, Moseley M, Ranum LP: **Non-ATG-initiated translation directed by microsatellite expansions.** *Proc Natl Acad Sci U S A* 2011, **108**:260–265.
73. Mattick JS, Makunin IV: **Small regulatory RNAs in mammals.** *Hum Mol Genet* 2005, **14**(1):R121–R132.
74. Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A, Tanzer A, Lagarde J, Lin W, Schlesinger F, Xue C, Marinov GK, Khatun J, Williams BA, Zaleski C, Rozowsky J, Röder M, Kokocinski F, Abdelhamid RF, Alioto T, Antoshechkin I, Baer MT, Bar NS, Batut P, Bell K, Bell I, Chakrabortty S, Chen X, Chrast J, Curado J, *et al*: **Landscape of transcription in human cells.** *Nature* 2012, **489**:101–108.
75. Yan B, Wang Z: **Long noncoding RNA: its physiological and pathological roles.** *DNA Cell Biol* 2012, **31**(Suppl 1):S34–S41.
76. Lee JT: **Epigenetic regulation by long noncoding RNAs.** *Science* 2012, **338**:1435–1439.
77. Faghihi MA, Wahlestedt C: **Regulatory roles of natural antisense transcripts.** *Nat Rev Mol Cell Biol* 2009, **10**:637–643.
78. Magistri M, Faghihi MA: **St Laurent G 3rd, Wahlestedt C: Regulation of chromatin structure by long noncoding RNAs: focus on natural antisense transcripts.** *Trends Genet* 2012, **28**:389–396.
79. Pastori C, Wahlestedt C: **Involvement of long noncoding RNAs in diseases affecting the central nervous system.** *RNA Biol* 2012, **9**:860–870.
80. Kumari D, Usdin K: **The distribution of repressive histone modifications on silenced FMR1 alleles provides clues to the mechanism of gene silencing in fragile X syndrome.** *Hum Mol Genet* 2010, **19**:4634–4642.
81. Ladd PD, Smith LE, Rabaia NA, Moore JM, Georges SA, Hansen RS, Hagerman RJ, Tassone F, Tapscott SJ, Filippova GN: **An antisense transcript spanning the CGG repeat region of FMR1 is upregulated in premutation carriers but silenced in full mutation individuals.** *Hum Mol Genet* 2007, **16**:3174–3187.
82. Khalil AM, Faghihi MA, Modarresi F, Brothers SP, Wahlestedt C: **A novel RNA transcript with antiapoptotic function is silenced in fragile X syndrome.** *PLoS One* 2008, **3**:e1486.
83. Pastori C, Peschansky VJ, Barbouth D, Mehta A, Silva JP, Wahlestedt C: **Comprehensive analysis of the transcriptional landscape of the human FMR1 gene reveals two new long noncoding RNAs differentially expressed in Fragile X syndrome and Fragile X-associated tremor/ataxia syndrome.** *Hum Genet* 2013.
84. Edbauer D, Neilson JR, Foster KA, Wang CF, Seeburg DP, Batterton MN, Tada T, Dolan BM, Sharp PA, Sheng M: **Regulation of synaptic structure and function by FMRP-associated microRNAs miR-125b and miR-132.** *Neuron* 2010, **65**:373–384.

doi:10.1186/1866-1955-6-23
Cite this article as: Sellier *et al.*: The multiple molecular facets of fragile X-associated tremor/ataxia syndrome. *Journal of Neurodevelopmental Disorders* 2014 **6**:23.

Chapter 6

The Molecular Biology of Premutation Expanded Alleles

Flora Tassone[1,2], Chantal Sellier[3], Nicolas Charlet-Berguerand[3], and Peter Todd[4,5]

[1] Department of Biochemistry and Molecular Medicine, School of Medicine, University of California, Davis, CA, USA.
[2] MIND Institute, University of California Davis Medical Center, Sacramento, CA, USA.
[3]IGBMC, CNRS UMR7104, INSERM U964, University of Strasbourg, Illkirch, France.
[4]Department of Neurology, University of Michigan, Ann Arbor, MI, USA.
[5]Staff Physician, VA Medical Center, Ann Arbor, MI, USA.
Email: ftassone@ucdavis.edu

Keywords: RAN translation, RNA disease, transcription, antisense, neurodegenerative disease, disease mechanisms

Abstract Fragile X-associated tremor/ataxia syndrome (FXTAS) is a late adult onset neurodegenerative disorder that mainly affects male carriers of an allele of 55- to 200 CGG repeats in the *FMR1* gene (premutation). FXTAS symptoms include progressive intention tremor, gait ataxia, neuropathy, psychiatric symptoms, cognitive decline, and autonomic dysfunctions. Neuropathological features of FXTAS include global cerebral and cerebellar atrophy, spongiform changes of white matter, marked Purkinje cell dropout and presence of ubiquitin-positive intranuclear inclusions throughout the brain. In contrast to fragile X syndrome, caused by transcriptional silencing of the *FMR1* gene FXTAS, is associated with elevated expression levels of the *FMR1* mRNA that although at lower levels, is also translated through non-canonical ATG translation into a polyGlycine containing protein (FMRpolyG). Here, we discuss current knowledge of the putative molecular mechanisms pathogenic to neuronal cell and underlying FXTAS.

Introduction

Although the developmental disorder, fragile X syndrome (FXS), is almost always caused by CGG repeat expansions exceeding 200 CGG repeats (full mutation range), with hypermethylation of the promoter region and consequent transcriptional silencing (Fu *et al.,* 1991; Oberlé *et al.,* 1991; Pieretti *et al.,* 1991; Verkerk *et al.,* 1991; Yu *et al.,* 1991), the adult onset neurodegenerative disorder, fragile X-associated tremor/ataxia syndrome (FXTAS), almost exclusively affects carriers of premutation alleles (55–200 CGG repeats) with nearly all cases having alleles that exceed ~65–70 CGG repeats

(Jacquemont *et al.,* 2006). Premutation alleles are generally unmethylated but several studies have reported on the presence of premutation alleles partial methylated even in the lower premutation size range and in some case the methylation status correlated with the severity of the observed clinical involvement (Allingham-Hawkins *et al.,* 1996; Pretto *et al.,* 2014c; Tassone *et al.,* 1999). Furthermore, several studies have reported FXTAS in individuals with a partially methylated full mutation (Loesch *et al.,* 2012; Pretto *et al.,* 2014a; Santa Maria *et al.,* 2013) and also in individuals with smaller alleles in the "gray zone" or intermediate range (45–54 CGG repeats; normal <45 CGG repeats) (Hall *et al.,* 2012; Liu *et al.,* 2013).

Concerning instability of the CGG repeat, alleles in the normal, gray and premutation size ranges generally have one or more AGG trinucleotides within the CGG repeat tract, typically spaced with intervals of 8–11 intervening CGG repeats (e.g., 9-10-10 pattern). However, these intervening repeats are generally lost with larger CGG repeat sizes within the premutation size range (Kunst *et al.,* 1994; Zhong *et al.,* 1996), which may result in the loss of repeat length stability (Zhong *et al.,* 1995). Interestingly, expansion from the premutation range to a full mutation allele occurs only during maternal transmission; that is, transmission of a full mutation allele never occurs from the father, even if the father is a carrier of a full mutation allele (Nolin *et al.,* 2003). Although the presence of AGG interruptions within the CGG repeat does not affect either the expression of the *FRM1* mRNA (Yrigollen *et al.,* 2012; Yrigollen *et al.,* 2014) or its translation (Ludwig *et al.,* 2009) it has been demonstrated to play a key role in the stability of the repeat (Nolin *et al.,* 2013; Yrigollen *et al.,* 2012; Yrigollen *et al.,* 2014). Importantly, Yigollen et al. demonstrated that the risk of expansion from premutation to full mutation alleles during maternal transmission decreases by increasing number of AGG interruptions. In addition, the risk of instability of an intermediate or premutation allele during maternal or paternal transmission is reduced with the presence of AGG interruptions, as is the magnitude of size change that occurs during transmission (Nolin *et al.,* 2013; Yrigollen *et al.,* 2014). Using logistic regression of 710 maternal transmissions and considering the total length of the CGG repeat allele, the number of AGG interruptions and the age of the mother at childbirth a model for measuring the risk of expansion to a full mutation during maternal transmission was calculated. This model was determined to be more suitable than that which considered pure CGG stretch instead of total length, confirming the previous data on a smaller sample (Yrigollen *et al.,* 2012).

The prevalence of full mutation is approximately 1 in 5,000 males and 1 in 2,500–8,000 females (reviewed in (Tassone *et al.,* 2012b)). However, based on a recent systematic review and meta-analysis, approximately 1/7000 males and 1/11000 females carry the full mutation (Hunter *et al.,* 2014).

Reported frequencies of the premutation allele vary among studies of various populations. Canadian studies reported a premutation frequency of approximately 1 in 250 for females and 1 in 800 for males (Dombrowski *et al.,* 2002; Rousseau *et al.,* 1995). However, the frequency of the premutation in females in Israel is substantially higher, at approximately 1 in 113 (Toledano-Alhadef *et al.,* 2001), and lower in an Asian population with a rate of 1 in 1,674 premutation males in Taiwan (Tzeng *et al.,* 2005). In

a recent report, using the known frequency for premutation females (using the average of 1 in 126 from (Pesso *et al.,* 2000; Toledano-Alhadef *et al.,* 2001)) to calculate the expected number of full mutations, Hagerman (Hagerman 2008) determined that the expected frequency for premutation males would be 1 in 282 and for full mutation (both males and females) would be 1 in 2,355. Interestingly, the former value is in agreement with what was observed in recent larger screening studies in which the prevalence of premutation alleles ranged between 150-210 in females and 290-430 in males (Hunter *et al.,* 2014; Maenner *et al.,* 2013; Tassone *et al.,* 2012b) .

Clinical Involvement in *FMR1* Premutation Carriers

Over the last several years, there has been increasing awareness of the spectrum of the phenotypes associated with premutation alleles of the *FMR1* gene (Hagerman *et al.,* 2013). While the presence of a full mutation allele (>200 CGG repeats) generally result in FXS, with autism spectrum disorders in over 60% of subjects (Harris *et al.,* 2008), smaller repeat expansions in the premutation range give rise to several distinct forms of clinical involvement: (a) fragile X-associated tremor/ataxia syndrome (FXTAS) mainly in older males; (b) Fragile X associated primary ovarian insufficiency FXPOI in approximately 20% of all carrier females; and may give rise to (c) behavioral, physical, emotional, and cognitive problems in some children who are premutation carriers, particularly at repeat sizes in the upper end of the premutation range.

FXTAS is mainly observed in males as demonstrated by the absence of significant differences between carrier females and controls when rated by CRST, ICARS or UPDRS (Berry-Kravis *et al.,* 2003; Jacquemont *et al.,* 2004), or FXTAS Rating Scale (Leehey *et al.,* 2007) – for details of use and findings with these rating scales in FXTAS the reader can see Chapter 1 and 7. However, a percentage of female carriers clearly present clinical and neuropathological signs of FXTAS (Berry-Kravis *et al.,* 2005; Jacquemont 2005; Peters *et al.,* 2006; Zuhlke *et al.,* 2004). In contrast to fragile X syndrome, both FXPOI and FXTAS appear to be generally confined to the premutation range with some exceptions (Liu *et al.,* 2013; Loesch *et al.,* 2012; Pretto *et al.,* 2014c; Santa Maria *et al.,* 2013). Moreover, while full mutations are characterized by the absence of *FMR1* mRNA, and consequently by the absence of FMRP, premutation carriers possess elevated levels of *FMR1* mRNA, with the extent of the increase depending upon the size of the premutation repeat expansion (Sellier *et al.,* 2014; Tassone *et al.,* 2000b). For the largest premutation alleles, FMRP levels are moderately reduced (up to twofold) due to a relative impediment to translation, demonstrated in both human and mouse (Ludwig *et al.,* 2014; Primerano *et al.,* 2002). A diagram of the relative levels of *FMR1* mRNA and proteins (FMRP) as a function of the number of CGG repeats and the association of clinical phenotypes is shown in Fig. 6.1.

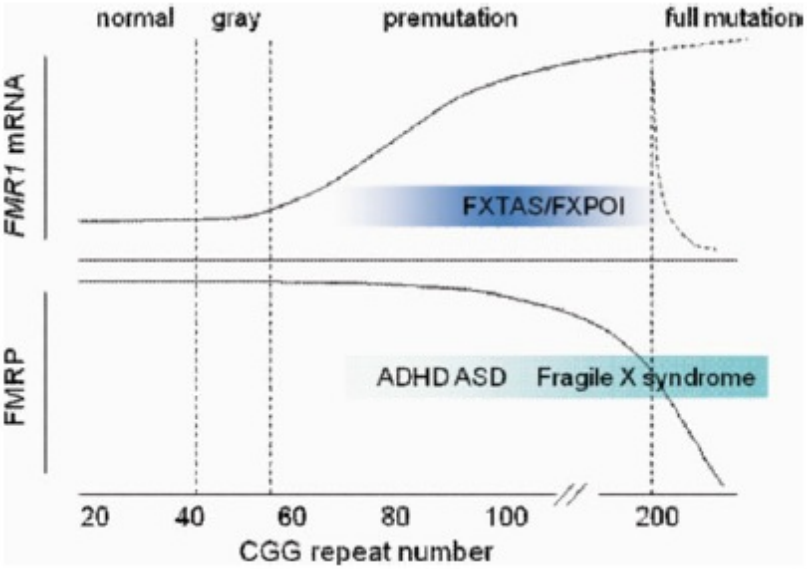


Fig. 6.1 FXTAS and FXPOI are largely confined to the premutation range and are thought to occur through an RNA toxic gain of function due to excess *FMR1* mRNA; however, patients with partially methylated full mutation alleles continue to express mRNA and are potentially at risk of developing FXTAS. By contrast, fragile X syndrome is caused by reduced/absent FMRP, due to silencing of the *FMR1* gene in the full mutation range. Features of the fragile X syndrome spectrum may also occur in the upper premutation range due to reduced protein production. *Dashed lines* for *FMR1* mRNA levels in the full mutation range reflect variations in degree of methylation; FMRP levels are reduced due to both lower mRNA levels and reduction in translation efficiency from the upper premutation range.

The expanded CGG repeat, located in the 5' UTR region of the *FMR1* gene, is now believed to cause the cellular toxicity in FXTAS, whereas FMRP is not thought to play a role in the neurodegenerative syndrome, since full mutation alleles (where FMRP is absent) are not associated with FXTAS. It is, however, possible that the decreased levels of FMRP could somehow act in concert with the effects of the toxic mRNA, particularly in the upper end of the premutation range where lower FMRP levels are detected; however, many cases of FXTAS are observed for CGG repeats in the 80–100 range, where FMRP levels are not substantially reduced. Thus, it appears clear that FXS and FXTAS are caused by a CGG expansion in the *FMR1* gene but with completely different molecular mechanisms (RNA Gain of Function for FXTAS; Loss of FMRP function for FXS).

As noted above, FXTAS is more frequent in males, as would be expected for an X-linked disorder (Adams *et al.,* 2007; Berry-Kravis *et al.,* 2007; Coffey S.M. *et al.,* 2008; Greco *et al.,* 2008; Hagerman *et al.,* 2004b; Hagerman *et al.,* 2007; Hagerman *et al.,* 2004c). In particular, female carriers are less likely to develop the disorder due to the protective effect of the expression of the normal allele in a portion of cells. Moreover, for those females who do experience features of FXTAS, expression of the normal *FMR1* allele is also likely to be responsible for the less severe clinical outcome (Hagerman *et al.,* 2004b; Zuhlke *et al.,* 2004) X-inactivation effects are therefore predicted to have a protective effect toward the neurological features of FXTAS. Several reports, in which a correlation between the severity of clinical signs and the X-inactivation ratio has been

observed, seem to be supportive of such theory (Berry-Kravis *et al.,* 2005; Coffey S.M. *et al.,* 2008; Jacquemont 2005). However, females can present with significant higher rates of muscle pain, thyroid problems, hypertension, and fibromyalgia (Coffey S.M. *et al.,* 2008; Hagerman *et al.,* 2013; Hundscheid *et al.,* 2003; Rodriguez-Revenga *et al.,* 2009).

Although details of the medical, cognitive, psychiatric, and neurological phenotypes of FXTAS and in premutation carriers in general, are discussed more extensively in other chapters of this book, a few additional comments regarding phenotype are pertinent here. Borderline to mild cognitive deficits generally are observed when FMRP levels are moderately reduced, as is often the case for individuals with alleles in the high premutation range (>150 CGG repeats) or for individuals with alleles in the low full mutation range that remain unmethylated and therefore retain transcriptional and somewhat translational activity (Tassone *et al.,* 2000b). Numerous phenotypic problems have been reported in premutation carriers that seem to begin in many cases in childhood. The most consistent deficits seen in such carriers of a large premutation include shyness, anxiety, social deficits, ADHD, and executive function deficits (Aziz *et al.,* 2003; Chonchaiya *et al.,* 2012; Cornish *et al.,* 2005; Farzin *et al.,* 2006). Thus, for the larger premutation alleles, cognitive involvement may reflect the combined effects of both RNA toxicity and lowered FMRP levels (Farzin *et al.,* 2006; Goodlin-Jones *et al.,* 2004; Tassone *et al.,* 2000a). The occurrence of seizures in about 8 to 13% of those with the premutation (Bailey Jr D.B. *et al.,* 2008; Chonchaiya *et al.,* 2012; Hagerman *et al.,* 2009), which is associated with the development of autism spectrum disorder (ASD) in males with the premutation (Chonchaiya *et al.,* 2012). Further, immune mediate disorders, including hypothyroidism and fibromyalgia, sleep apnea, hypertension and migraine are more common in premutation carriers than controls (Hagerman *et al.,* 2013), Finally, very recent studies of early development in infant premutation carriers has evidentiated the presence processing deficits similar to infants with FXS. Specifically, a visual discrimination deficit in babies with the premutation that is similar to that seen in babies with the full mutation and is significantly different from visual discrimination in age- and sex-matched control babies has been recently reported and suggesting that spatiotemporal processing impairment may constitute an endophenotype in infant premutation carriers (Gallego *et al.,* 2014). In addition, examination of cognitive, communication, and social-behavioral profiles has demonstrated subtle developmental differences in premutation infants as young as 12 months of age (Wheeler *et al.,* 2015).

A number of studies have reported on the correlation between the length of the CGG repeat within the premutation range and the severity of clinical involvement in premutation carriers. These correlations are particularly strong with the neurological (FXTAS) phenotypes including the age of onset of tremor and ataxia (Tassone *et al.,* 2007a), overall motor impairment (Leehey *et al.,* 2007), severity of white matter disease and degree of brain atrophy (Cohen *et al.,* 2006; Loesch *et al.,* 2005), severity of neuropathic signs (Berry-Kravis *et al.,* 2007), degree of neuropathy as measured by nerve conduction studies (Soontarapornchai *et al.,* 2008), reduced cerebellar volume (Adams *et al.,* 2007), and the percent of inclusions and age at death (Greco *et al.,* 2006). These associations are discussed further in Chapter 5.

FMR1 Gene Structure

The *FMR1* gene consists of 17 exons spanning 38 kb of Xq27.3 (Eichler *et al.,* 1993). The gene is expressed in many tissues; however, the highest expression of the 4.4 kb transcript is observed in brain, placenta, testis, lung, and kidney (Hinds *et al.,* 1993). In both human, including fetal brain neurons, and mouse, extensive alternative splicing of the *FMR1* gene has been demonstrated by RT-PCR analysis, and several FMRP isoforms have been observed on Westerns (Ashley *et al.,* 1993; Huang *et al.,* 1996; Sittler 1996; Verheij 1993; Verkerk 1993) due to alternative splicing in the carboxy-terminal half of the *FMR1* gene. Published studies have suggested that alternative splicing in the *FMR1* gene does not seem to be tissue specific, as similar ratios of transcripts were found in several fetal tissues, including brain and testis (Verkerk 1993); however, quantitative approaches were not taken. The main splice variants observed in the *FMR1* gene involve the use of alternative splice acceptors in exons 12, 14, 15, and 17 (Ashley CT 1993; Verkerk 1993). The existence of 16 different mRNA isoforms has been recently described in different tissues (Pretto *et al.,* 2015). Although the relative abundance of these isoforms was reported to be significantly increased in premutations, interestingly the abundance of isoforms spliced at both exons 12 and 14, specifically *Iso10 and Iso10b,* containing the complete exon 15 and differing only in splicing in exon 17 was 4-6 fold increased compared to controls. Based on their findings the authors suggested that RNA toxicity is likely to arise from an increased level of all *FMR1* mRNA isoforms and such increase may have a functional relevance in the pathology of *FMR1* associated disorders. Although the specific function of the splice variants is not known, four or five different FMRP isoforms have been described using FMRP-specific antibodies (Verheij *et al.,* 1995) some of which act as shuttling proteins that transports their mRNA targets from the nucleus to the cytoplasm (Dury *et al.,* 2013).

In addition to the presence of alternative splicing of the *FMR1* gene, a detailed analysis of the CG-rich, TATA-less, promoter region of the *FMR1* gene has revealed an influence of the CGG repeat with respect to initiation or the start of initiation. The existence of three distinct groups of transcriptional initiation sites, and the distribution of these start sites, which is modulated by the number of CGG repeats in the downstream (5' UTR) region, has been demonstrated in different tissues including lymphocytes and neuronal cells (Beilina *et al.,* 2004; Tassone *et al.,* 2011). While premutation alleles appear to preferentially express the longer *FMR1* mRNA (50 bp longer compared to the most frequent transcription initiation site in the normal alleles) (site II or site III), normal alleles appear to preferentially use the shorter transcripts initiating at site I (Beilina *et al.,* 2004; Tassone *et al.,* 2011). Interestingly, the nucleotide sequence of all three transcriptional initiation sites was found to be highly similar to the consensus sequence of pyrimidine-rich initiator (Inr) elements (consensus sequence YYAN(T/A)YY) (Javahery *et al.,* 1994), which are usually located near the start site and have been implicated in transcription initiation in TATA-less genes (Chow *et al.,* 1995). In addition, a fourth transcription initiation site, located between the previously identified sites I and II, was identified in brain tissues, suggesting the presence of a brain-specific transcription start site (Tassone *et al.,* 2011). The dependency of alternative

transcription initiation site usage on the CGG repeats length has also been observed in mice (Tassone *et al.,* 2011).

Alternative polyadenylation site usage in the 3'UTR, which is implicated in the regulation of gene expression of many genes, has also been observed in the *FMR1* mRNA. Several different mRNA transcripts through the usage of different polyadenylation signals were identified in both human and mouse brain tissue (Tassone *et al.,* 2011). Thus, the combination of both alternative 5**'** and 3**'** UTRs in addition the complexity of expression of the *FMR1* isoforms suggest their potential role in neuronal physiology, as well as in *FMR1* associated disorders; however it needs to be investigated.

Importantly, the differential usage of initiators in normal and premutation alleles may imply a possible variation in translational efficiency (post-transcriptional regulation) due to variation in 5' structure and sequence. Indeed, the FMRP deficit observed, especially in the upper premutation range, is CGG dependent and is due to decreased translational efficiency. Reduced translational efficiency was observed both in cell lines and in transient transfection experiments using expanded alleles spanning the entire premutation range (Chen *et al.,* 2003; Primerano *et al.,* 2002). Similar findings are observed in two independently generated knockin mouse models of FXTAS (Brouwer *et al.,* 2008; Entezam *et al.,* 2007; Iliff *et al.,* 2013; Ludwig *et al.,* 2014; Willemsen *et al.,* 2003). To date, the precise mechanism by which the expanded CGG repeat impedes translation is not clear. Translation from the larger premutation alleles is expected to be severely inhibited by the predicted free energies of stabilization of the CGG repeat element. One interesting feature of the transition from normal to the premutation range is the gradual loss of AGG interruptions within the CGG element located in the 5' UTR region of the *FMR1* mRNA.

As mentioned earlier, a few studies have shown that both transcription and translation expression levels (Ludwig *et al.,* 2009; Tassone *et al.,* 2007b; Yrigollen *et al.,* 2012) of the *FMR1* mRNA are not influenced by AGG interruptions. Therefore, even if the presence of AGG repeats may modify the secondary/tertiary structure of the UTR region of the *FMR1* gene, such alterations do not seem to have an effect on transcription or represent a rate-limiting step in translational initiation.

Higher *FMR1* Transcription Rate in Premutation Alleles

While hypermethylation of the *FMR1* gene in full mutation alleles leads to transcriptional silencing, premutation alleles are generally unmethylated and, therefore, transcriptionally active. Moreover, *FMR1* message levels are 2- to 10-fold higher than normal, particularly in the upper end of the premutation range (~100–200 CGG repeats), despite their reduced protein (FMRP) levels (Sellier *et al.,* 2014). Both elevated mRNA levels (up to fivefold) and FMRP deficit were also observed in some fragile X males with a partially methylated full mutation, even for alleles greater than 300 CGG repeats (Tassone *et al.,* 2000b).

Elevated *FMR1* mRNA levels observed in premutation carriers appear to result not from increased mRNA stability (Tassone *et al.,* 2000b) but rather from increased transcriptional activity of the *FMR1* gene. This was best demonstrated by increased

levels of run-on transcription in premutation alleles compared to normal (Tassone *et al.,* 2007a). Nuclear retention of expanded repeat *FMR1* mRNA was not observed by RNA in situ hybridization experiments in patient derived lymphocytes (Tassone *et al.,* 2007a), although this has not yet been evaluated in neurons. Elevated levels of *FMR1* messenger RNA and a deficit in translation efficiency in expanded alleles has been evaluated using *in vitro* translation experiments with a luciferase reporter mRNA containing the 5'UTR of the *FMR1* gene containing varying numbers of CGG repeats. Specifically, the translation efficiency gradually decreases with increasing CGG repeat number (Chen *et al.,* 2003).

The augmented transcription of *FMR1* mRNA is associated with epigenetic alterations induced by the CGG repeat expansion itself (Todd *et al.,* 2010). Specifically, ectopically expressed CGG repeat expansions are sufficient to elicit chromatin changes in a *Drosophila* model of the disease and similar changes are observed in patient derived cell lines. Interestingly, these alterations in local chromatin structure are dynamic and modifiable by genetic and pharmacologic means, suggesting that agents aimed at chromatin remodeling might have therapeutic potential in FXTAS (Todd *et al.,* 2010). How exactly the repeat lead to these epigenetic alterations is unclear. Long tracts of CGG-repeats are known to exclude nucleosomes in vitro (Wang *et al.,* 1996). Should this also occur in vivo it could lead to enhanced transcription by enhancing access of transcription factors to the promoter. Alternatively, recent reports demonstrate that transcribed CGG-Repeats form R-loop structures that could also potentially trigger increased *FMR1* expression (Groh *et al.,* 2014; Loomis *et al.,* 2014; Usdin *et al.,* 2014). Future work will be required to define the sequential pathway and proximal events that lead to transcriptional up-regulation.

***ASFMR1*: The Antisense Transcript at the *FMR1* Locus**

Using RACE analysis and strand-specific RT-PCR, Ladd and collaborators (Ladd *et al.,* 2007) identified an antisense transcript of the *FMR1* gene *(ASFMR1)* overlapping with the CGG repeat region. The *ASFMR1* is widely expressed in human tissues, with a higher expression in brain and similarly to the *FMR1* gene, the expression appears to be influenced by the CGG repeat number as it is upregulated in peripheral blood leukocytes of premutation carriers and silenced in full mutation alleles. Moreover, *ASFMR1* transcript is alternatively spliced, polyadenylated, exported to the cytoplasm and contains an open reading frame encompassing the CCG repeat that potentially encodes a polyproline peptide (Ladd *et al.,* 2007). Shortly after this report, Khalil and collaborators described a second antisense transcript to *FMR1*, *FMR4*, which originates upstream of the *FMR1* start site and covers 2.4 kb of sequence (Khalil *et al.,* 2008). Expression of *FMR4*, like that of *ASFMR1* and *FMR1*, is increased in brain from premutation individuals and silenced in individuals with the full mutation (Khalil *et al.,* 2008). Importantly, *FMR4* over-expression increases cell proliferation while *FMR4* down-regulation induces apoptosis in vitro (Khalil *et al.,* 2008). More recently, two novel transcripts arising from the *FMR1* locus, *FMR5* and *FMR6*, were identified (Pastori *et al.,* 2013). *FMR5* is a sense-oriented long non-coding RNA (lncRNA) transcribed from

approximately 1 kb upstream of the *FMR1* transcription start site. In contrast to *FMR1*, the *FMR5* transcript is not differentially expressed in human brain from unaffected individuals compared to full and premutation patients, suggesting that its transcription is independent of CGG repeat expansion (Pastori *et al.,* 2013). *FMR6* is a spliced 600 nt long antisense transcript whose sequence is complementary to the 3' region of *FMR1*. *FMR6* begins in the 3'UTR, ends in exon 15 of *FMR1*, and uses the same splice junctions as *FMR1*. Unexpectedly, the expression of *FMR6* is reduced in premutation carriers suggesting that abnormal transcription and/or chromatin remodeling occurs toward the distal end of the locus. However, the chromatin marks associated with the 3' end of *FMR1* in premutation carriers are yet to be described and further studies are needed to determine the potential contribution of these long non coding RNA to the variable clinical phenotypes associated with FXTAS and to the *FMR1* associated disorders.

RNA Toxic Gain-of-Function Model

As stated above, FXTAS has been observed, with rare exceptions, only in premutation carriers (Loesch *et al.,* 2012; Pretto *et al.,* 2014c; Santa Maria *et al.,* 2013). The absence of FXTAS in carriers of fully silenced *FMR1* alleles has led to the hypothesis that FXTAS is the result of a toxic gain-of-function of the *FMR1* mRNA itself (Hagerman *et al.,* 2004a; Hagerman *et al.,* 2001). This hypothesis is based on a precedent set by the myotonic dystrophies, DM1 and DM2, in which expression of mutant RNAs containing hundreds to thousands of CUG (DM1) or CCUG (DM2) repeats accumulate in nuclear RNA aggregates that sequester specific RNA binding proteins (the Muscleblind-like proteins). Depletion of the free pool of MBNL proteins leads to specific RNA metabolisms changes, which ultimately results in the symptoms of myotonic dystrophy (Nelson *et al.,* 2013; Ranum *et al.,* 2006).

Extending this RNA gain-of-function model to FXTAS predicts that the mutant *FMR1* RNA containing expanded CGG repeats may sequester specific proteins resulting in loss of their normal functions, which would ultimately cause the symptoms of FXTAS (Hagerman *et al.,* 2004a). Consistent with this hypothesis, Iwahashi and collaborators identified more than 20 proteins from inclusions purified from brains of FXTAS patients (Iwahashi *et al.,* 2006). Among these, two RNA binding proteins were of special interest, hnRNP A2/B1, that is mutated in families with inherited degeneration affecting muscle, brain, bone and motor neurons (Kim *et al.,* 2013), and MBNL1, the RNA binding protein that is sequestered in myotonic dystrophy (Kanadia *et al.,* 2003; Miller *et al.,* 2000). However, a role for MBNL1 in FXTAS appears unlikely, since no genetic interaction between MBNL1 and CGG-mediated neurodegeneration was observed in fly model of FXTAS (Sofola *et al.,* 2007), and no misregulation of splicing events regulated by MBNL1 were observed in brain samples from FXTAS patients (Sellier *et al.,* 2010).

Association of hnRNP A2/B1 to RNA containing expanded CGG repeats was confirmed by several independent analyses (Jin *et al.,* 2007; Sellier *et al.,* 2010; Sofola *et al.,* 2007). Specifically, the interaction of hnRNP A2/B1 with RNA containing expanded CGG repeats was observed in cytoplasmic cerebellar lysates (Sofola *et al.,* 2007), and the

importance of the titration of the cytoplasmic pool of hnRNP A2/B1 by expanded CGG repeats was demonstrated by the impaired dendritic delivery of the BC1 RNA, a known target of hnRNP A2/B1 (Muslimov *et al.,* 2011). In contrast, nuclear hnRNP A2/B1 exibited little binding to CGG RNA (Sofola *et al.,* 2007), and alternative splicing regulated by nuclear hnRNP A2/B1 were not altered in brain samples of FXTAS patients (Sellier *et al.,* 2010). These data suggest that some modifications of hnRNP A2/B1, either in the nucleus or in the cytoplasm, may modify the ability of hnRNP A2/B1 to bind to CGG RNA, so that expanded CGG repeats may recruit and deplete the cytoplasmic pool of hnRNP A2/B1, but not the nuclear pool of hnRNP A2/B1. In addition, the ability of hnRNP A proteins to unfold RNA structures formed by expanded CGG repeats (Khateb *et al.,* 2004; Ofer *et al.,* 2009), raises the interesting hypothesis that hnRNP A2/B1 may also act as a RNA chaperone that destabilizes the RNA structures formed by expanded CGG repeats.

hnRNP A2/B1 recruits other proteins, such as the CUGBP1 RNA binding protein, to RNA containing expanded CGG repeats (Sofola *et al.,* 2007). Overexpression of either hnRNP A2/B1 or CUGBP1 rescues the neurodegeneration in a *Drosophila* expressing 90 CGG repeats, highlighting the potential importance of hnRNP A2/B1 and CUGBP1 to FXTAS pathology (Jin *et al.,* 2007; Sofola *et al.,* 2007). Similarly, hnRNP A2/B1 interacts with the heterochromatin protein 1 (HP1) to silence the expression of specific *Drosophila* retrotransposons, such as *gypsy* or *copia*, and the titration of hnRNP A2/B1 by expanded CGG repeats results in increased expression of these retrotransposons in fly model of FXTAS (Tan *et al.,* 2012). Knockdown of *gypsy* RNA expression, but not of *copia*, suppress the toxicity induced by expanded CGG repeats in *Drosophila (Tan et al., 2012)*. Whether the expression of retrotransposons is altered in FXTAS patients remains to be determined, but it is interesting to note that increased levels of 5 - hydroxymethylcytosine (5hmc) have been identified in repetitive elements DNA in a mouse model of FXTAS (Yao *et al.,* 2014).

In addition to hnRNP A2/B1, proteomic analyses revealed that SAM68, a splicing regulator encoded by the *KHDRBS1* gene, was also found in CGG RNA aggregates (Sellier *et al.,* 2010). However, overexpression of SAM68 was not sufficient to rescue neuronal cell death induced by expression of expanded CGG repeats. Similarly to hnRNP A2/B1 that recruits CUGBP1 through protein-protein interactions, SAM68 did not bind directly to CGG repeats and recruitment of SAM68 within CGG RNA aggregates occurred in *trans*.

Sellier and collaborators (2013) found that DROSHA-DGCR8, the enzymatic complex that processes pri-microRNAs into pre-miRNAs, associates directly with expanded CGG repeats. Sequestration of DROSHA-DGCR8 within CGG RNA aggregates resulted in reduced processing of pri-miRNAs in neuronal cells expressing expanded CGG repeats, and overexpression of DGCR8 rescued neuronal cell death induced by expression of CGG repeats (Sellier *et al.,* 2013). DROSHA-DGCR8 also recruits SAM68 to CGG repeats. Other studies also demonstrated alteration of the expression of specific miRNAs in blood samples of FXTAS patients and in *Drosophila* expressing CGG repeat (Alvarez-Mora *et al.,* 2013; Tan *et al.,* 2012). These data suggest that misregulation of the processing and/or of the expression of miRNA may be of importance for the pathogenicity of FXTAS.

Proteomic analysis performed by Jin and collaborators (2007) showed that Purα (purine-rich binding protein α) binds to RNA containing expanded CGG repeats. Purα is a single-stranded cytoplasmic DNA- and RNA-binding protein that has been implicated in many biological processes, including RNA transport and translation. Importantly, overexpression of Purα rescued neurodegeneration in a *Drosophila* model of FXTAS (Jin *et al.,* 2007). However, the presence of Purα within nuclear aggregates in FXTAS brain samples was not consistently observed. Jin et al. (2007) found Purα in cytoplasmic inclusions in *Drosophila* expressing 90 CGG repeats. In contrast, Purα was not identified within the intranuclear inclusions in brain sections of mouse premutation models, or in hippocampal or cortical brain sections derived from patients with FXTAS (Galloway *et al.,* 2009; Iwahashi *et al.,* 2006; Sellier *et al.,* 2013). These results indicate that composition of the inclusions may vary from one model organism to the other.

As with hnRNP A2/B1 that recruits CUGBP1, Purα recruits Rm62, the *Drosophila* ortholog of the RNA helicase P68/DDX5 (Qurashi *et al.,* 2011). Overexpression of Rm62 rescued neurodegeneration in flies expressing 90 CGG repeats, highlighting the potential importance of P68/DDX5 to FXTAS pathology (Qurashi *et al.,* 2011). Of interest, the RNA helicases P68/DDX5 and DDX6 have been recently reported to be involved in the unfolding of expanded CUG repeats in myotonic dystrophy (Laurent *et al.,* 2012; Pettersson *et al.,* 2014). These data raise the hypothesis that regulating the RNA structures of CUG expanded repeats in DM, or of CGG repeats in FXTAS may be of therapeutic interest (Disney *et al.,* 2012; Tran *et al.,* 2014; Yang *et al.,* 2015).

Finally, simultaneous studies by Peter Todd and David Nelson's groups demonstrated that overexpression of TDP-43, encoded by the *TARDBP* gene, suppresses the neurodegeneration induced by expanded CGG repeats in *Drosophila* (Galloway *et al.,* 2009; He *et al.,* 2014). Interestingly, the expression of *Tardbp* is down-regulated in Purkinje neurons of a mouse model of FXTAS expressing the 5'UTR of *FMR1* with 90 CGG repeats under the Pcp2 promoter (Galloway *et al.,* 2009). TDP-43 is of special interest for neurodegenerative diseases since protein inclusions of TDP-43 in neurons is a histopathological hallmark of amyotrophic lateral sclerosis (ALS) (Neumann *et al.,* 2006), and mutation of *TARDBP* leads to amyotrophic lateral sclerosis and frontotemporal dementia (ALS-FTD) (Kunst *et al.,* 1994). Interestingly, TDP-43 is not found within the RNA aggregates of CGG repeats, but requires the hnRNP A2/B1 homologues Hrb87F and Hrb98DE to modulate the CGG repeat triggered toxicity (He *et al.,* 2014). Overall, these observations suggest that mutant *FMR1* RNA containing expanded CGG repeats could be pathogenic by sequestering specific RNA binding proteins, resulting in loss of their normal functions, ultimately leading to neuronal cell dysfunction and death. In that aspect, it remains to be determined whether expression of Purα, hnRNP A2/B1, P68/DDX5, DROSHA-DGCR8, CUGBP1 or TDP-43 rescues any phenotype in mouse models expressing expanded CGG repeats. Such functional rescue experiments will be instrumental to determine the importance of these candidate RNA binding proteins to FXTAS pathology.

Inclusion Formation

The presence of eosinophilic intranuclear inclusions, broadly distributed throughout the brain and brainstem of affected individuals (Greco *et al.,* 2006; Greco *et al.,* 2002; Tassone *et al.,* 2004) represent the neuropathological hallmark of FXTAS. These inclusions are present as single, spherical (~2–5 μm of diameter) particles that are found exclusively within the nuclei. Inclusions have been identified in both neurons and astrocytes throughout the cerebrum and the brain stem, including cells of the ependymal and sub-ependymal layers (Tassone *et al.,* 2004). They are most numerous in the hippocampus and are rarely seen in Purkinje cells of the cerebellum (Greco *et al.,* 2006; Greco *et al.,* 2002). Outside of the CNS, rare inclusions have been observed in both the pineal and the posterior and anterior pituitary glands but also in the pancreas, in the heart, in the thyroid and adrenal glands, in the Leydig cells of the testis, in ganglia and in periadrenal fat tissue (Greco *et al.,* 2007; Hunsaker *et al.,* 2011). Inclusions have also been observed in postmortem brain tissue of females with FXTAS (Hagerman *et al.,* 2004c; Tassone *et al.,* 2012a).

Pioneering work from Hagerman and Tassone's laboratories demonstrated that these inclusions are ubiquitin-positive and contain various chaperone proteins such as Hsp27, Hsp70 and αB-crystallin, but are negative for tau proteins, α-synuclein, or polyglutamine (Greco *et al.,* 2002; Iwahashi *et al.,* 2006). Importantly, these inclusions also stain positive for the *FMR1* RNA messenger as well as for its polyGlycine RAN-translated product (Buijsen *et al.,* 2014; Tassone *et al.,* 2004; Todd *et al.,* 2010) suggesting that expression of mutant *FMR1* mRNA may trigger inclusion formation. In support of this hypothesis, the formation of inclusions can be recapitulated in neuronal cell cultures using a minimal construct expressing an expansion of 99 CGG repeats embedded within the 5'UTR of *FMR1* (Arocena *et al.,* 2005; Sellier *et al.,* 2010; Todd *et al.,* 2010). Of interest, plasmids harboring either no expanded CGG repeats or expanded CGG repeat but under an inactive promoter (no RNA) did not lead to inclusion formation. Also, intranuclear inclusions formed even when the portion encoding FMRP was deleted (Arocena *et al.,* 2005; Todd *et al.,* 2010). Thus, expression of the 5'UTR of *FMR1* containing a CGG repeat expansion is necessary and sufficient for the formation of these inclusions.

These observations in neuronal cell cultures are consistent with the work of Jin et al. (2003), who showed that expression in *Drosophila* of a construct containing the 5'UTR of *FMR1* with 90 CGG repeat induces the formation of cytoplasmic ubiquitin-positive inclusions associated with neurodegeneration and eye pathology (Jin *et al.,* 2003). Similarly, a knock-in (KI) mouse model, in which the endogenous eight CGG repeats of the murine *Fmr1* were replaced with an expansion containing ~100 CGG repeats of human origin, demonstrated ubiquitin-positive nuclear inclusions and mild neuromotor and behavioral disturbances (Brouwer *et al.,* 2008; Van Dam *et al.,* 2005; Willemsen *et al.,* 2003). A result confirmed by recent mouse models where heterologous expression of the sole 5'UTR of *FMR1* containing expanded CGG repeats under strong chimeric promoters is sufficient to cause cellular toxicity, despite the presence of normal *Fmr1* alleles and unaltered expression of Fmrp (Hashem *et al.,* 2009; Hukema *et al.,* 2014).

These animal models demonstrate that the expression of *FMR1* mRNA containing expanded CGG repeats is both necessary and sufficient to cause the pathological features of human FXTAS (Berman *et al.,* 2014).

RAN translation produces a polyglycine-containing protein in FXTAS.

What drives inclusion formation in FXTAS was unclear until recently. The working hypothesis in the field was that the CGG RNA repeats were serving as a nidus for inclusion formation by nucleating together a set of RNA binding proteins. However, while some inclusions in brain samples of FXTAS patients contain mutant *FMR1* RNA with expanded CGG repeats (Tassone *et al.,* 2004), mouse models in which express expanded CGG repeats show numerous ubiquitin-positive inclusions but only rarely RNA aggregates, suggesting that other factors might trigger inclusion formation (Sellier *et al.,* 2013). Moreover, the inclusions observed in FXTAS are ubiquitin positive, which is a characteristic not commonly observed in other RNA-dominant disorders such as Myotonic Dystrophy.

A hint for how nucleotide repeats in the 5'UTR of a messenger RNA might trigger inclusion formation came with the discovery of Repeat associated non-AUG (RAN) translation by Laura Ranum and colleagues (Zu *et al.,* 2011). They found that expression of either CAG in cells outside the context of an AUG-initiated open reading frame were capable of supporting translational initiation and production of homopolymeric proteins. Surprisingly, this RAN translation occurred in all three possible reading frames to produce polyglutamine, polyalanine, and polyserine containing proteins. Initiation was largely independent of the surrounding sequence context but was influenced by the cell type in which the repeats were expressed and was somewhat different for different repeat frames. They then established that these alternative "RAN" products accumulated in tissues in animal models of Spinocerebellar Ataxia type 8 and in lymphoblasts from patients with DM1 (where a CAG repeat is expressed as an antisense transcript). Since these initial observations, RAN translation has been reported in association with CUG repeats (as seen in DM1) and with both GGGGCC and CCCCGG repeat expansions in C9orf72 that cause ALS and Frontotemporal Dementia (Ash *et al.,* 2013; Gendron *et al.,* 2013; Mori *et al.,* 2013a; Mori *et al.,* 2013b; Zu *et al.,* 2011; Zu *et al.,* 2013).

Todd and colleagues (2013) independently discovered that placement of CGG repeats upstream of a fluorescent reporter protein (GFP) in certain reading frames led to aggregation of GFP in transfected cells and in *Drosophila*. This aggregation was associated with production of a higher molecular weight species that resulted from translational initiation within the 5'UTR just above the CGG repeat. They were able to demonstrate by mass spectroscopy and biochemical analysis that this unconventional translational initiation (CGG RAN translation) occurred at one of several "near" AUG codons (one nucleotide different from AUG) to produce products in two reading frames, producing a polyglycine containing protein (FMRpolyG) and a polyalanine containing protein (FMRpolyA) (Todd *et al.,* 2013). Production of FMRpolyG was confirmed in patients by use of antibodies generated against the predicted C terminus of this novel

protein. The FMRpolyG protein was found in a fraction of ubiquitinated inclusions in brain autopsy samples from patients (Todd *et al.,* 2013). Recent studies with a second set of antibodies covering different N-terminal and C-terminal epitopes revealed that most of the ubiquitin-positive inclusions observed in FXTAS brain sections are also FMRpolyG positive (Buijsen *et al.,* 2014).

A role for FMRpolyG in disease pathogenesis is supported by work in *Drosophila.* In both transfected cells and in flies, it was demonstrated that moving the repeat to the 3'UTR or placing a stop codon just above the repeat prevented FMRpolyG production and suppressed CGG repeat associated toxicity. In contrast, placing an AUG start codon above the repeat enhanced FMRpolyG production and led to greater lethality in flies and cells. However, the mechanism by which FMRpolyG induces toxicity is still unclear. Impairment of the ubiquitin proteasome system (UPS) enhances CGG repeat associated toxicity (Handa *et al.,* 2005; Oh *et al.,* 2015; Todd *et al.,* 2013) and expression of chaperone proteins such as HSP-70 suppresses CGG repeat associated toxicity in flies (Jin *et al.,* 2003). This genetic interaction between UPS impairment and CGG toxicity is largely driven by RAN translation (Oh *et al.,* 2015). Moreover, RAN translation is capable of triggering UPS impairment in transfected cells (Oh *et al.,* 2015). Together these studies suggest that failures in protein quality control are likely one method by which RAN translation cause toxicity, but future work will be needed to define how this unusual process contributes to FXTAS pathogenesis.

Disruption of LaminA/C architecture.

Among the various proteins that Iwahashi and collaborators identified within the inclusions in brain samples of FXTAS patients (Iwahashi *et al.,* 2006), the lamin A/C proteins are of special interest. Lamin proteins constitute a matrix of protein that is located next to the inner nuclear membrane and that is involved in nuclear stability, chromatin structure and gene expression. Mutations in the *LMNA* gene, which encodes both lamin A and C splicing forms, are associated with several diseases called laminopathies, including limb girdle and Emery-Dreyfuss muscular dystrophies, dilated cardiomyopathy, lipodystrophy, Charcot-Marie-Tooth disease and Hutchinson-Gilford progeria syndrome. Interestingly, nuclear organization of lamin A/C is disrupted in MEF of knock-in mice expressing ~200 CGG repeats, as well as in skin fibroblasts of FXTAS patients with loss of the typical architecture of lamin A/C within the nuclear ring (Garcia-Arocena *et al.,* 2010). These alterations of the lamin A/C structure can be recapitulated in cell cultures upon transfection of a minimal construct expressing expanded CGG repeats embedded within the 5'UTR of *FMR1* (Arocena *et al.,* 2005; Hoem *et al.,* 2011; Sellier *et al.,* 2010). Also, the expression of *Lmna* mRNA is decreased in Purkinje neurons of a mouse model of FXTAS expressing 90 CGG repeats under the Pcp2 promoter (Galloway *et al.,* 2014). In contrast, the expression of *LMNA* mRNA is increased while the amounts of soluble protein of Lamin A and C are slightly reduced in brain samples of FXTAS patients (Garcia-Arocena *et al.,* 2010). Overall, these results suggest that lamin A/C architecture and expression are altered in FXTAS, however it remains to determine the molecular mechanism leading to these alterations. It is therefore intriguing that

recent data on GGGGCC repeat expansions in *C9orf72* which cause ALS and FTD are associated with changes in nuclear envelope architecture and nuclear pore complex function (Freibaum *et al.,* 2015; Jovicic *et al.,* 2015; Zhang *et al.,* 2015). Thus, the pathological consequences of lamin alterations in FXTAS are an exciting area of research in need of further exploration.

Mitochondrial dysfunction

Mitochondrial biogenesis plays a central role in neurogenesis and neuroplasticity and several studies have demonstrated significant mitochondrial dysfunction in premutation carriers (Kaplan *et al.,* 2012; Napoli *et al.,* 2011; Ross-Inta *et al.,* 2010), both with and without fragile X-associated tremor ataxia syndrome (FXTAS). Specifically, altered biochemical characteristics pointing to a lower ATP production, including decreased NAD- and FAD-linked oxygen uptake rates, decreased mitochondrial protein expression, increased oxidative/nitrative stress and altered zinc bioavailability were demonstrated in both cultured dermal fibroblasts and brain samples from premutation carriers with and without FXTAS, within a wide range of CGG repeats (Napoli *et al.,* 2011; Ross-Inta *et al.,* 2010). Kaplan et al. (2012) observed fewer mitochondria and greatly reduced mitochondria mobility in hippocampal neuronal culture from preCGG KI mice, expressing low FMRP levels and higher *Fmr1* mRNA than that measured in wild type during the early stages of development. In addition, preCGG neurons presented with abnormal metabolic function including higher rates of basal oxygen consumption, ATP production and proton leak. The authors suggested that deficits in mitochondrial trafficking and metabolic function could play a role in the early pathophysiology observed in premutation carriers and also potentially contribute to the risk of developing FXTAS.

In addition, an altered Ca2+ regulation was reported in both neuronal cultures from the premutation mouse model (*Fmr1* preCGG mouse) (Cao *et al.,* 2012) but also in iPSC derived from premutation carriers (Liu *et al.,* 2012). The impaired Ca2+ regulation was associated with alteration of the most abundant excitatory neurotransmitter in the vertebrate CNS, the Glutammate, with a decrease in the expression of GLT-1 and GLAST Glutamate uptake and of their expression levels in astrocytes cultures from preCGG mice and in postmortem cerebellum of PM carriers with FXTAS compared with age matched controls. Higher CGG repeat number had the greatest reductions in both proteins (Cao *et al.,* 2013; Pretto *et al.,* 2014b). Abnormal clustered burst (CB) firing electrical activity and abnormal patterns of synchronized calcium oscillations in the cytosol were also observed with the premutation mouse neurons, which, interestingly, were reversed with mGluR5 antagonists or the GABAA receptor positive modulator allopregnanolone (Cao *et al.,* 2012) making it a potential candidate for a beneficial therapeutic approach in FXTAS.

Finally, shorter telomeres have been observed in a number of neurodegenerative conditions including Alzheimer disease, cell senescence, and also in Down syndrome (Jenkins *et al.,* 2006; Panossian *et al.,* 2003). The role of telomeres includes protection against degeneration; they also appear to be essential for chromosomal stability and could lead to cell senescence. Telomere shortening has also been observed in a small

sample of male carriers of the premutation (Jenkins *et al.,* 2008) with FXTAS and with or without dementia. However, further studies are warranted to establish if telomere shortening could be considered a biomarker for the cellular dysfunction observed in FXTAS.

The above observations suggest that many triggering events (Figure 2) can be associated with CGG expansion leading to cellular dysregulation and dysfunction and ultimately, potentially, to neurodegeneration and *FMR1* associated disorders.

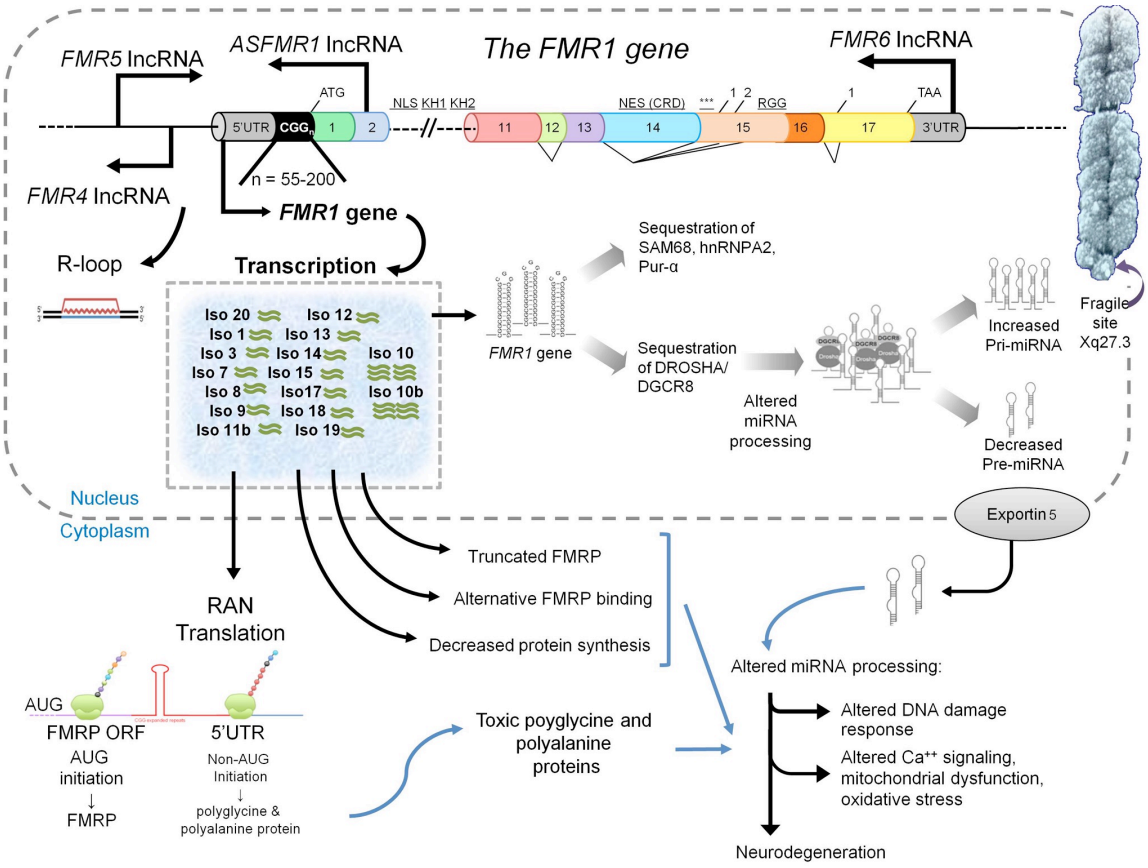


Fig. 6.2 Diagram showing the molecular alterations of premutation expanded alleles including increased *FMR1* mRNA expression levels, R-Loop formation, sequestration of CGG binding proteins, miRNA dysregulation and RAN translation.

REFERENCES

Adams, J. S., Adams, P. E., Nguyen, D., Brunberg, J. A., Tassone, F., Zhang, W., Koldewyn, K., Rivera, S. M., Grigsby, J., Zhang, L., Decarli, C., Hagerman, P. J., Hagerman, R. J. (2007). Volumetric brain changes in females with fragile X-associated tremor/ataxia syndrome (FXTAS). Neurology 69(9):851-9.

Allingham-Hawkins, D. J., Brown, C. A., Babul, R., Chitayat, D., Krekewich, K., Humphries, T., Ray, P. N., Teshima, I. E. (1996). Tissue-specific methylation differences and cognitive function in fragile X premutation females. American Journal of Medical Genetics 64(2):329-33.

Alvarez-Mora, M. I., Rodriguez-Revenga, L., Madrigal, I., Torres-Silva, F., Mateu-Huertas, E., Lizano, E., Friedlander, M. R., Marti, E., Estivill, X., Mila, M. (2013). MicroRNA expression profiling in blood from fragile X-associated tremor/ataxia syndrome patients. Genes Brain Behav 12(6):595-603.

Arocena, D. G., Iwahashi, C. K., Won, N., Beilina, A., Ludwig, A. L., Tassone, F., Schwartz, P. H., Hagerman, P. J. (2005). Induction of inclusion formation and disruption of lamin A/C structure by premutation CGG-repeat RNA in human cultured neural cells. Hum Mol Genet 14(23):3661-3671.

Ash, P. E., Bieniek, K. F., Gendron, T. F., Caulfield, T., Lin, W. L., Dejesus-Hernandez, M., Van Blitterswijk, M. M., Jansen-West, K., Paul, J. W., 3rd, Rademakers, R., Boylan, K. B., Dickson, D. W., Petrucelli, L. (2013). Unconventional translation of C9ORF72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. Neuron 77(4):639-46.

Ashley, C. T., Jr., Wilkinson, K. D., Reines, D., Warren, S. T. (1993). FMR1 protein: conserved RNP family domains and selective RNA binding. Science 262(5133):563-6.

Ashley Ct, S. J., Kunst Cb, Leiner Ha, Eichler Ee, Nelson Dl, Warren St. (1993). Human and murine FMR-1: alternative splicing and translational initiation downstream of the CGG-repeat. Nat Genet 4(3):244-51.

Aziz, M., Stathopulu, E., Callias, M., Taylor, C., Turk, J., Oostra, B., Willemsen, R., Patton, M. (2003). Clinical features of boys with fragile X premutations and intermediate alleles. Am J Med Genet 121B(1):119-27.

Bailey Jr D.B., Raspa M., Olmsted M., D.B., H. (2008). Co-Occurring Conditions Associated With FMR1 Gene Variations: Findings From a National Parent Survey. Am J Med Genet 146A(16):2060-2069.

Beilina, A., Tassone, F., Schwartz, P. H., Sahota, P., Hagerman, P. J. (2004). Redistribution of transcription start sites within the FMR1 promoter region with expansion of the downstream CGG-repeat element. Hum Mol Genet 13(5):543-9.

Berman, R. F., Buijsen, R. A., Usdin, K., Pintado, E., Kooy, F., Pretto, D., Pessah, I. N., Nelson, D. L., Zalewski, Z., Charlet-Bergeurand, N., Willemsen, R., Hukema, R. K. (2014). Mouse models of the fragile X premutation and fragile X-associated tremor/ataxia syndrome. J Neurodev Disord 6(1):25.

Berry-Kravis, E., Abrams, L., Coffey, S. M., Hall, D. A., Greco, C., Gane, L. W., Grigsby, J., Bourgeois, J. A., Finucane, B., Jacquemont, S., Brunberg, J. A., Zhang, L., Lin, J., Tassone, F., Hagerman, P. J., Hagerman, R. J., Leehey, M. A. (2007). Fragile X-associated tremor/ataxia syndrome: Clinical features, genetics, and testing guidelines. Mov Disord 22(14):2018-30.

Berry-Kravis, E., Lewin, F., Wuu, J., Leehey, M., Hagerman, R., Hagerman, P., Goetz, C. G. (2003). Tremor and ataxia in fragile X premutation carriers: blinded videotape study. Ann Neurol 53(5):616-23.

Berry-Kravis, E., Potanos, K., Weinberg, D., Zhou, L., Goetz, C. G. (2005). Fragile X-associated tremor/ataxia syndrome in sisters related to X-inactivation. Ann Neurol 57(1):144-147.

Brouwer, J. R., Severijnen, E., De Jong, F. H., Hessl, D., Hagerman, R. J., Oostra, B. A., Willemsen, R. (2008). Altered hypothalamus-pituitary-adrenal gland axis regulation in the expanded CGG-repeat mouse model for fragile X-associated tremor/ataxia syndrome. Psychoneuroendocrinology 33(6):863-73.

Buijsen, R. A., Sellier, C., Severijnen, L. A., Oulad-Abdelghani, M., Verhagen, R. F., Berman, R. F., Charlet-Berguerand, N., Willemsen, R., Hukema, R. K. (2014). FMRpolyG-positive inclusions in CNS and non-CNS organs of a fragile X premutation carrier with fragile X-associated tremor/ataxia syndrome. Acta Neuropathol Commun 2:162.

Cao, Z., Hulsizer, S., Cui, Y., Pretto, D. L., Kim, K. H., Hagerman, P. J., Tassone, F., Pessah, I. N. (2013). Enhanced asynchronous Ca(2+) oscillations associated with impaired glutamate transport in cortical astrocytes expressing Fmr1 gene premutation expansion. J Biol Chem 288(19):13831-41.

Cao, Z., Hulsizer, S., Tassone, F., Tang, H. T., Hagerman, R. J., Rogawski, M. A., Hagerman, P. J., Pessah, I. N. (2012). Clustered burst firing in FMR1 premutation hippocampal neurons: amelioration with allopregnanolone. Hum Mol Genet 21(13):2923-35.

Chen, L. S., Tassone, F., Sahota, P., Hagerman, P. J. (2003). The (CGG)n repeat element within the 5' untranslated region of the FMR1 message provides both positive and negative cis effects on in vivo translation of a downstream reporter. Hum Mol Genet 12(23):3067-74.

Chonchaiya, W., Au, J., Schneider, A., Hessl, D., Harris, S. W., Laird, M., Mu, Y., Tassone, F., Nguyen, D. V., Hagerman, R. J. (2012). Increased prevalence of seizures in boys who were probands with the FMR1 premutation and co-morbid autism spectrum disorder. Hum Genet 131(4):581-9.

Chow, C. W., Clark, M. P., Rinaldo, J. E., Chalkley, R. (1995). Multiple initiators and C/EBP binding sites are involved in transcription from the TATA-less rat XDH/XO basal promoter. Nucleic Acids Res 23(16):3132-40.

Coffey S.M., Cook K., Tartaglia N., Tassone F., Nguyen D.V., Pan R., Bronsky H.E., Yuhas J., Borodyanskaya M., Grigsby J., Doerflinger M., Hagerman P.J., R.J., H. (2008). Expanded clinical phenotype of women with the FMR1 premutation. Am J Med Genet A 146A(8):1009-16.

Cohen, S., Masyn, K., Adams, J., Hessl, D., Rivera, S., Tassone, F., Brunberg, J., Decarli, C., Zhang, L., Cogswell, J., Loesch, D., Leehey, M., Grigsby, J., Hagerman, P. J., Hagerman, R. J. (2006). Molecular and imaging correlates of the fragile X-associated tremor/ataxia syndrome. Neurology 67(8):1426-31.

Cornish, K., Kogan, C., Turk, J., Manly, T., James, N., Mills, A., Dalton, A. (2005). The emerging fragile X premutation phenotype: evidence from the domain of social cognition. Brain Cogn 57(1):53-60.

Disney, M. D., Liu, B., Yang, W. Y., Sellier, C., Tran, T., Charlet-Berguerand, N., Childs-Disney, J. L. (2012). A small molecule that targets r(CGG)(exp) and improves defects in fragile X-associated tremor ataxia syndrome. ACS Chem Biol 7(10):1711-8.

Dombrowski, C., Levesque, M. L., Morel, M. L., Rouillard, P., Morgan, K., Rousseau, F. (2002). Premutation and intermediate-size FMR1 alleles in 10 572 males from the general population: loss of an AGG interruption is a late event in the generation of fragile X syndrome alleles. Hum Mol Genet 11(4):371-378.

Dury, A. Y., El Fatimy, R., Tremblay, S., Rose, T. M., Cote, J., De Koninck, P., Khandjian, E. W. (2013). Nuclear Fragile X Mental Retardation Protein is localized to Cajal bodies. PLoS Genet 9(10):e1003890.

Eichler, E. E., Richards, S., Gibbs, R. A., Nelson, D. L. (1993). Fine structure of the human FMR1 gene [published erratum appears in Hum Mol Genet 1994 Apr;3(4):684-5]. Hum Mol Genet 2(8):1147-53.

Entezam, A., Biacsi, R., Orrison, B., Saha, T., Hoffman, G. E., Grabczyk, E., Nussbaum, R. L., Usdin, K. (2007). Regional FMRP deficits and large repeat expansions into the full mutation range in a new Fragile X premutation mouse model. Gene 395(1-2):125-34.

Farzin, F., Perry, H., Hessl, D., Loesch, D., Cohen, J., Bacalman, S., Gane, L., Tassone, F., Hagerman, P., Hagerman, R. (2006). Autism spectrum disorders and attention-deficit/hyperactivity disorder in boys with the fragile X premutation. J Dev Behav Pediatr 27(2 Suppl):S137-44.

Freibaum, B. D., Lu, Y., Lopez-Gonzalez, R., Kim, N. C., Almeida, S., Lee, K.-H., Badders, N., Valentine, M., Miller, B. L., Wong, P. C. (2015). GGGGCC repeat expansion in C9orf72 compromises nucleocytoplasmic transport. Nature.

Fu, Y. H., Kuhl, D. P., Pizzuti, A., Pieretti, M., Sutcliffe, J. S., Richards, S., Verkerk, A. J., Holden, J. J., Fenwick, R. G., Jr., Warren, S. T., Oostra, B. A., Nelson, D. L., Caskey, C. T. (1991). Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. Cell 67(6):1047-58.

Gallego, P., Burris, J., Rivera, S. (2014). Visual motion processing deficits in infants with the fragile X premutation. J Neurodev Disord 6:29.

Galloway, J. N., Nelson, D. L. (2009). Evidence for RNA-mediated toxicity in the fragile X-associated tremor/ataxia syndrome. Future Neurol 4(6):785.

Galloway, J. N., Shaw, C., Yu, P., Parghi, D., Poidevin, M., Jin, P., Nelson, D. L. (2014). CGG repeats in RNA modulate expression of TDP-43 in mouse and fly models of fragile X tremor ataxia syndrome. Hum Mol Genet 23(22):5906-15.

Garcia-Arocena, D., Yang, J. E., Brouwer, J. R., Tassone, F., Iwahashi, C., Berry-Kravis, E. M., Goetz, C. G., Sumis, A. M., Zhou, L., Nguyen, D. V., Campos, L., Howell, E., Ludwig, A., Greco, C., Willemsen, R., Hagerman, R. J., Hagerman, P. J. (2010). Fibroblast phenotype in male carriers of FMR1 premutation alleles. Hum Mol Genet 19(2):299-312.

Gendron, T. F., Cosio, D. M., Petrucelli, L. (2013). c9RAN translation: a potential therapeutic target for the treatment of amyotrophic lateral sclerosis and frontotemporal dementia. Expert Opin Ther Targets 17(9):991-5.

Goodlin-Jones, B. L., Tassone, F., Gane, L. W., Hagerman, R. J. (2004). Autistic spectrum disorder and the fragile X premutation. J Dev Behav Pediatr 25(6):392-8.

Greco, C. M., Berman, R. F., Martin, R. M., Tassone, F., Schwartz, P. H., Chang, A., Trapp, B. D., Iwahashi, C., Brunberg, J., Grigsby, J., Hessl, D., Becker, E. J., Papazian, J., Leehey, M. A., Hagerman, R. J., Hagerman, P. J. (2006). Neuropathology of fragile X-associated tremor/ataxia syndrome (FXTAS). Brain 129(Pt 1):243-55.

Greco, C. M., Hagerman, R. J., Tassone, F., Chudley, A. E., Del Bigio, M. R., Jacquemont, S., Leehey, M., Hagerman, P. J. (2002). Neuronal intranuclear inclusions in a new cerebellar tremor/ataxia syndrome among fragile X carriers. Brain 125(Pt 8):1760-71.

Greco, C. M., Soontarapornchai, K., Wirojanan, J., Gould, J. E., Hagerman, P. J., Hagerman, R. J. (2007). Testicular and pituitary inclusion formation in fragile X associated tremor/ataxia syndrome. J Urol 177(4):1434-7.

Greco, C. M., Tassone, F., Garcia-Arocena, D., Tartaglia, N., Coffey, S. M., Vartanian, T. K., Brunberg, J. A., Hagerman, P. J., Hagerman, R. J. (2008). Clinical and neuropathologic

findings in a woman with the FMR1 premutation and multiple sclerosis. Arch Neurol 65(8):1114-6.

Groh, M., Lufino, M. M., Wade-Martins, R., Gromak, N. (2014). R-loops associated with triplet repeat expansions promote gene silencing in Friedreich ataxia and fragile X syndrome. PLoS Genet 10(5):e1004318.

Hagerman, P. J. (2008). The fragile X prevalence paradox. J Med Genet 45(8):498-9.

Hagerman, P. J., Hagerman, R. J. (2004a). The fragile-X premutation: a maturing perspective. Am J Hum Genet 74(5):805-16.

Hagerman, P. J., Hagerman, R. J. (2004b). Fragile X-associated Tremor/Ataxia Syndrome (FXTAS). Ment Retard Dev Disabil Res Rev 10(1):25-30.

Hagerman, P. J., Stafstrom, C. E. (2009). Origins of epilepsy in fragile X syndrome. Epilepsy Curr 9(4):108-12.

Hagerman, R., Hagerman, P. (2013). Advances in clinical and molecular understanding of the FMR1 premutation and fragile X-associated tremor/ataxia syndrome. Lancet Neurol 12(8):786-98.

Hagerman, R. J., Coffey, S. M., Maselli, R., Soontarapornchai, K., Brunberg, J. A., Leehey, M. A., Zhang, L., Gane, L. W., Fenton-Farrell, G., Tassone, F., Hagerman, P. J. (2007). Neuropathy as a presenting feature in fragile X-associated tremor/ataxia syndrome. Am J Med Genet A 143(19):2256-60.

Hagerman, R. J., Leavitt, B. R., Farzin, F., Jacquemont, S., Greco, C. M., Brunberg, J. A., Tassone, F., Hessl, D., Harris, S. W., Zhang, L., Jardini, T., Gane, L. W., Ferranti, J., Ruiz, L., Leehey, M. A., Grigsby, J., Hagerman, P. J. (2004c). Fragile-X-associated tremor/ataxia syndrome (FXTAS) in females with the FMR1 premutation. Am J Hum Genet 74(5):1051-6.

Hagerman, R. J., Leehey, M., Heinrichs, W., Tassone, F., Wilson, R., Hills, J., Grigsby, J., Gage, B., Hagerman, P. J. (2001). Intention tremor, parkinsonism, and generalized brain atrophy in male carriers of fragile X. Neurology 57:127-130.

Hall, D., Tassone, F., Klepitskaya, O., Leehey, M. (2012). Fragile X–associated tremor ataxia syndrome in FMR1 gray zone allele carriers. Mov Disord 27(2):297-301.

Handa, V., Goldwater, D., Stiles, D., Cam, M., Poy, G., Kumari, D., Usdin, K. (2005). Long CGG-repeat tracts are toxic to human cells: implications for carriers of Fragile X premutation alleles. FEBS Lett 579(12):2702-8.

Harris, S. W., Hessl, D., Goodlin-Jones, B., Ferranti, J., Bacalman, S., Barbato, I., Tassone, F., Hagerman, P. J., Herman, H., Hagerman, R. J. (2008). Autism profiles of males with fragile X syndrome. American Journal of Mental Retardation 113(6):427-38.

Hashem, V., Galloway, J. N., Mori, M., Willemsen, R., Oostra, B. A., Paylor, R., Nelson, D. L. (2009). Ectopic expression of CGG containing mRNA is neurotoxic in mammals. Hum Mol Genet 18(13):2443-51.

He, F., Krans, A., Freibaum, B. D., Taylor, J. P., Todd, P. K. (2014). TDP-43 suppresses CGG repeat-induced neurotoxicity through interactions with HnRNP A2/B1. Hum Mol Genet 23(19):5036-51.

Hinds, H. L., Ashley, C. T., Sutcliffe, J. S., Nelson, D. L., Warren, S. T., Housman, D. E., Schalling, M. (1993). Tissue specific expression of FMR-1 provides evidence for a functional role in fragile X syndrome [see comments] [published erratum appears in Nat Genet 1993 Nov;5(3):312]. Nat Genet 3(1):36-43.

Hoem, G., Raske, C. R., Garcia-Arocena, D., Tassone, F., Sanchez, E., Ludwig, A. L., Iwahashi, C. K., Kumar, M., Yang, J. E., Hagerman, P. J. (2011). CGG-repeat length threshold for FMR1 RNA pathogenesis in a cellular model for FXTAS. Hum Mol Genet 20(11):2161-70.

Huang, T., Li, L. Y., Shen, Y., Qin, X. B., Pang, Z. L., Wu, G. Y. (1996). Alternative splicing of the FMR1 gene in human fetal brain neurons. Am J Med Genet 64(2):252-5.

Hukema, R. K., Riemslagh, F. W., Melhem, S., Van Der Linde, H. C., Severijnen, L. A., Edbauer, D., Maas, A., Charlet-Berguerand, N., Willemsen, R., Van Swieten, J. C. (2014). A new inducible transgenic mouse model for C9orf72-associated GGGGCC repeat expansion supports a gain-of-function mechanism in C9orf72-associated ALS and FTD. Acta Neuropathol Commun 2:166.

Hundscheid, R. D., Smits, A. P., Thomas, C. M., Kiemeney, L. A., Braat, D. D. (2003). Female carriers of fragile X premutations have no increased risk for additional diseases other than premature ovarian failure. Am J Med Genet A 117(1):6-9.

Hunsaker, M. R., Greco, C. M., Spath, M. A., Smits, A. P., Navarro, C. S., Tassone, F., Kros, J. M., Severijnen, L. A., Berry-Kravis, E. M., Berman, R. F., Hagerman, P. J., Willemsen, R., Hagerman, R. J., Hukema, R. K. (2011). Widespread non-central nervous system organ pathology in fragile X premutation carriers with fragile X-associated tremor/ataxia syndrome and CGG knock-in mice. Acta Neuropathol 122(4):467-79.

Hunter, J., Rivero-Arias, O., Angelov, A., Kim, E., Fotheringham, I., Leal, J. (2014). Epidemiology of fragile X syndrome: a systematic review and meta-analysis. Am J Med Genet A 164A(7):1648-58.

Iliff, A. J., Renoux, A. J., Krans, A., Usdin, K., Sutton, M. A., Todd, P. K. (2013). Impaired activity-dependent FMRP translation and enhanced mGluR-dependent LTD in Fragile X premutation mice. Hum Mol Genet 22(6):1180-92.

Iwahashi, C. K., Yasui, D. H., An, H. J., Greco, C. M., Tassone, F., Nannen, K., Babineau, B., Lebrilla, C. B., Hagerman, R. J., Hagerman, P. J. (2006). Protein composition of the intranuclear inclusions of FXTAS. Brain 129(Pt 1):256-71.

Jacquemont, S. (2005). Screening for FXTAS. Eur J Hum Genet 13(1):2-3.

Jacquemont, S., Farzin, F., Hall, D., Leehey, M., Tassone, F., Gane, L., Zhang, L., Grigsby, J., Jardini, T., Lewin, F., Berry-Kravis, E., Hagerman, P. J., Hagerman, R. J. (2004). Aging in individuals with the FMR1 mutation. Am J Ment Retard 109(2):154-64.

Jacquemont, S., Leehey, M. A., Hagerman, R. J., Beckett, L. A., Hagerman, P. J. (2006). Size bias of fragile X premutation alleles in late-onset movement disorders. J Med Genet 43(10):804-9.

Javahery, R., Khachi, A., Lo, K., Zenzie-Gregory, B., Smale, S. T. (1994). DNA sequence requirements for transcriptional initiator activity in mammalian cells. Mol Cell Biol 14(1):116-27.

Jenkins, E. C., Velinov, M. T., Ye, L., Gu, H., Li, S., Jenkins, E. C., Jr., Brooks, S. S., Pang, D., Devenny, D. A., Zigman, W. B., Schupf, N., Silverman, W. P. (2006). Telomere shortening in T lymphocytes of older individuals with Down syndrome and dementia. Neurobiol Aging 27(7):941-5.

Jenkins, E. C., Ye, L., Gu, H., Ni, S. A., Duncan, C. J., Velinov, M., Pang, D., Krinsky-Mchale, S. J., Zigman, W. B., Schupf, N., Silverman, W. P. (2008). Increased "absence" of telomeres may indicate Alzheimer's disease/dementia status in older individuals with Down syndrome. Neurosci Lett 440(3):340-3.

Jin, P., Duan, R., Qurashi, A., Qin, Y., Tian, D., Rosser, T. C., Liu, H., Feng, Y., Warren, S. T. (2007). Pur alpha binds to rCGG repeats and modulates repeat-mediated neurodegeneration in a Drosophila model of fragile X tremor/ataxia syndrome. Neuron 55(4):556-64.

Jin, P., Zarnescu, D. C., Zhang, F., Pearson, C. E., Lucchesi, J. C., Moses, K., Warren, S. T. (2003). RNA-mediated neurodegeneration caused by the fragile X premutation rCGG repeats in Drosophila. Neuron 39(5):739-47.

Jovicic, A., Mertens, J., Boeynaems, S., Bogaert, E., Chai, N., Yamada, S. B., Paul, J. W., 3rd, Sun, S., Herdy, J. R., Bieri, G., Kramer, N. J., Gage, F. H., Van Den Bosch, L., Robberecht, W., Gitler, A. D. (2015). Modifiers of C9orf72 dipeptide repeat toxicity connect nucleocytoplasmic transport defects to FTD/ALS. Nat Neurosci 18(9):1226-9.

Kanadia, R. N., Johnstone, K. A., Mankodi, A., Lungu, C., Thornton, C. A., Esson, D., Timmers, A. M., Hauswirth, W. W., Swanson, M. S. (2003). A muscleblind knockout model for myotonic dystrophy. Science 302(5652):1978-80.

Kaplan, E. S., Cao, Z., Hulsizer, S., Tassone, F., Berman, R. F., Hagerman, P. J., Pessah, I. N. (2012). Early mitochondrial abnormalities in hippocampal neurons cultured from Fmr1 pre-mutation mouse model. J Neurochem 123(4):613-21.

Khalil, A. M., Faghihi, M. A., Modarresi, F., Brothers, S. P., Wahlestedt, C. (2008). A novel RNA transcript with antiapoptotic function is silenced in fragile X syndrome. PLoS One 3(1):e1486.

Khateb, S., Weisman-Shomer, P., Hershco, I., Loeb, L. A., Fry, M. (2004). Destabilization of tetraplex structures of the fragile X repeat sequence (CGG)n is mediated by homolog-conserved domains in three members of the hnRNP family. Nucleic Acids Res 32(14):4145-54.

Kim, H. J., Kim, N. C., Wang, Y. D., Scarborough, E. A., Moore, J., Diaz, Z., Maclea, K. S., Freibaum, B., Li, S., Molliex, A., Kanagaraj, A. P., Carter, R., Boylan, K. B., Wojtas, A. M., Rademakers, R., Pinkus, J. L., Greenberg, S. A., Trojanowski, J. Q., Traynor, B. J., Smith, B. N., Topp, S., Gkazi, A. S., Miller, J., Shaw, C. E., Kottlors, M., Kirschner, J., Pestronk, A., Li, Y. R., Ford, A. F., Gitler, A. D., Benatar, M., King, O. D., Kimonis, V. E., Ross, E. D., Weihl, C. C., Shorter, J., Taylor, J. P. (2013). Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. Nature 495(7442):467-73.

Kunst, C. B., Warren, S. T. (1994). Cryptic and polar variation of the fragile X repeat could result in predisposing normal alleles. Cell 77(6):853-61.

Ladd, P. D., Smith, L. E., Rabaia, N. A., Moore, J. M., Georges, S. A., Hansen, R. S., Hagerman, R. J., Tassone, F., Tapscott, S. J., Filippova, G. N. (2007). An antisense transcript spanning the CGG repeat region of FMR1 is upregulated in premutation carriers but silenced in full mutation individuals. Hum Mol Genet 16(24):3174-87.

Laurent, F. X., Sureau, A., Klein, A. F., Trouslard, F., Gasnier, E., Furling, D., Marie, J. (2012). New function for the RNA helicase p68/DDX5 as a modifier of MBNL1 activity on expanded CUG repeats. Nucleic Acids Res 40(7):3159-71.

Leehey, M. A., Berry-Kravis, E., Min, S. J., Hall, D. A., Rice, C. D., Zhang, L., Grigsby, J., Greco, C. M., Reynolds, A., Lara, R., Cogswell, J., Jacquemont, S., Hessl, D. R., Tassone, F., Hagerman, R., Hagerman, P. J. (2007). Progression of tremor and ataxia in male carriers of the FMR1 premutation. Mov Disord 22(2):203-6.

Liu, J., Koscielska, K. A., Cao, Z., Hulsizer, S., Grace, N., Mitchell, G., Nacey, C., Githinji, J., Mcgee, J., Garcia-Arocena, D., Hagerman, R. J., Nolta, J., Pessah, I. N., Hagerman, P. J. (2012). Signaling defects in iPSC-derived fragile X premutation neurons. Hum Mol Genet 21(17):3795-805.

Liu, Y., Winarni, T., Zhang, L., Tassone, F., Hagerman, R. (2013). Fragile X-associated tremor/ataxia syndrome (FXTAS) in grey zone carriers. Clin Genet 84(1):74-77.

Loesch, D. Z., Litewka, L., Brotchie, P., Huggins, R. M., Tassone, F., Cook, M. (2005). Magnetic resonance imaging study in older fragile X premutation male carriers. Ann Neurol 58(2):326-330.

Loesch, D. Z., Sherwell, S., Kinsella, G., Tassone, F., Taylor, A., Amor, D., Sung, S., Evans, A. (2012). Fragile X-associated tremor/ataxia phenotype in a male carrier of unmethylated full mutation in the FMR1 gene. Clin Genet 82(1):88-92.

Loomis, E. W., Sanz, L. A., Chedin, F., Hagerman, P. J. (2014). Transcription-associated R-loop formation across the human FMR1 CGG-repeat region. PLoS Genet 10(4):e1004294.

Ludwig, A. L., Espinal, G. M., Pretto, D. I., Jamal, A. L., Arque, G., Tassone, F., Berman, R. F., Hagerman, P. J. (2014). CNS expression of murine fragile X protein (FMRP) as a function of CGG-repeat size. Hum Mol Genet 23(12):3228-38.

Ludwig, A. L., Raske, C., Tassone, F., Garcia-Arocena, D., Hershey, J. W., Hagerman, P. J. (2009). Translation of the FMR1 mRNA is not influenced by AGG interruptions. Nucleic Acids Res 37(20):6896-904.

Maenner, M. J., Baker, M. W., Broman, K. W., Tian, J., Barnes, J. K., Atkins, A., Mcpherson, E., Hong, J., Brilliant, M. H., Mailick, M. R. (2013). FMR1 CGG expansions: prevalence and sex ratios. Am J Med Genet B Neuropsychiatr Genet 162B(5):466-73.

Miller, J. W., Urbinati, C. R., Teng-Umnuay, P., Stenberg, M. G., Byrne, B. J., Thornton, C. A., Swanson, M. S. (2000). Recruitment of human muscleblind proteins to (CUG)(n) expansions associated with myotonic dystrophy. EMBO J 19(17):4439-48.

Mori, K., Arzberger, T., Grasser, F. A., Gijselinck, I., May, S., Rentzsch, K., Weng, S. M., Schludi, M. H., Van Der Zee, J., Cruts, M., Van Broeckhoven, C., Kremmer, E., Kretzschmar, H. A., Haass, C., Edbauer, D. (2013a). Bidirectional transcripts of the expanded C9orf72 hexanucleotide repeat are translated into aggregating dipeptide repeat proteins. Acta Neuropathol 126(6):881-93.

Mori, K., Weng, S. M., Arzberger, T., May, S., Rentzsch, K., Kremmer, E., Schmid, B., Kretzschmar, H. A., Cruts, M., Van Broeckhoven, C., Haass, C., Edbauer, D. (2013b). The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTLD/ALS. Science 339(6125):1335-8.

Muslimov, I. A., Patel, M. V., Rose, A., Tiedge, H. (2011). Spatial code recognition in neuronal RNA targeting: role of RNA-hnRNP A2 interactions. J Cell Biol 194(3):441-57.

Napoli, E., Ross-Inta, C., Wong, S., Omanska-Klusek, A., Barrow, C., Iwahashi, C., Garcia-Arocena, D., Sakaguchi, D., Berry-Kravis, E., Hagerman, R., Hagerman, P. J., Giulivi, C. (2011). Altered zinc transport disrupts mitochondrial protein processing/import in fragile X-associated tremor/ataxia syndrome. Hum Mol Genet 20(15):3079-92.

Nelson, D. L., Orr, H. T., Warren, S. T. (2013). The unstable repeats--three evolving faces of neurological disease. Neuron 77(5):825-43.

Neumann, M., Sampathu, D. M., Kwong, L. K., Truax, A. C., Micsenyi, M. C., Chou, T. T., Bruce, J., Schuck, T., Grossman, M., Clark, C. M., Mccluskey, L. F., Miller, B. L., Masliah, E., Mackenzie, I. R., Feldman, H., Feiden, W., Kretzschmar, H. A., Trojanowski, J. Q., Lee, V. M. (2006). Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Science 314(5796):130-3.

Nolin, S. L., Brown, W. T., Glicksman, A., Houck, G. E., Gargano, A. D., Sullivan, A., Biancalana, V., Brondum-Nielsen, K., Hjalgrim, H., Holinski-Feder, E., Kooy, F., Longshore, J., Macpherson, J., Mandel, J. L., Matthijs, G., Rousseau, F., Steinbach, P., Vaisanen, M. L., Von Koskull, H., Sherman, S. (2003). Expansion of the fragile X CGG repeat in females with premutation or intermediate alleles. Am J Hum Genet 72:454-464.

Nolin, S. L., Sah, S., Glicksman, A., Sherman, S. L., Allen, E., Berry-Kravis, E., Tassone, F., Yrigollen, C., Cronister, A., Jodah, M., Ersalesi, N., Dobkin, C., Brown, W. T., Shroff, R., Latham, G. J., Hadd, A. G. (2013). Fragile X AGG analysis provides new risk predictions for 45-69 repeat alleles. Am J Med Genet A 161(4):771-8.

Oberlé, I., Rousseau, F., Heitz, D., Kretz, C., Devys, D., Hanauer, A., Boue, J., Bertheas, M. F., Mandel, J. L. (1991). Instability of a 550-base pair DNA segment and abnormal methylation in fragile X syndrome. Science 252(5010):1097-102.

Ofer, N., Weisman-Shomer, P., Shklover, J., Fry, M. (2009). The quadruplex r(CGG)n destabilizing cationic porphyrin TMPyP4 cooperates with hnRNPs to increase the translation efficiency of fragile X premutation mRNA. Nucleic Acids Res 37(8):2712-22.

Oh, S. Y., He, F., Krans, A., Frazer, M., Taylor, J. P., Paulson, H. L., Todd, P. K. (2015). RAN translation at CGG repeats induces ubiquitin proteasome system impairment in models of fragile X-associated tremor ataxia syndrome. Hum Mol Genet 24(15):4317-26.

Panossian, L. A., Porter, V. R., Valenzuela, H. F., Zhu, X., Reback, E., Masterman, D., Cummings, J. L., Effros, R. B. (2003). Telomere shortening in T cells correlates with Alzheimer's disease status. Neurobiol Aging 24(1):77-84.

Pastori, C., Peschansky, V. J., Barbouth, D., Mehta, A., Silva, J. P., Wahlestedt, C. (2013). Comprehensive analysis of the transcriptional landscape of the human FMR1 gene reveals two new long noncoding RNAs differentially expressed in Fragile X syndrome and Fragile X-associated tremor/ataxia syndrome. Hum Genet.

Pesso, R., Berkenstadt, M., Cuckle, H., Gak, E., Peleg, L., Frydman, M., Barkai, G. (2000). Screening for fragile X syndrome in women of reproductive age. Prenat Diagn 20(8):611-14.

Peters, N., Kamm, C., Asmus, F., Holinski-Feder, E., Kraft, E., Dichgans, M., Bruning, R., Gasser, T., Botzel, K. (2006). Intrafamilial variability in fragile X-associated tremor/ataxia syndrome. Movement Disorders 21:98-102.

Pettersson, O. J., Aagaard, L., Andrejeva, D., Thomsen, R., Jensen, T. G., Damgaard, C. K. (2014). DDX6 regulates sequestered nuclear CUG-expanded DMPK-mRNA in dystrophia myotonica type 1. Nucleic Acids Res 42(11):7186-200.

Pieretti, M., Zhang, F. P., Fu, Y. H., Warren, S. T., Oostra, B. A., Caskey, C. T., Nelson, D. L. (1991). Absence of expression of the FMR-1 gene in fragile X syndrome. Cell 66(4):817-22.

Pretto, D., Yrigollen, C. M., Tang, H. T., Williamson, J., Espinal, G., Iwahashi, C. K., Durbin-Johnson, B., Hagerman, R. J., Hagerman, P. J., Tassone, F. (2014a). Clinical and molecular implications of mosaicism in FMR1 full mutations. Front Genet 5:318.

Pretto, D. I., Eid, J. S., Yrigollen, C. M., Tang, H. T., Loomis, E. W., Raske, C., Durbin-Johnson, B., Hagerman, P. J., Tassone, F. (2015). Differential increases of specific FMR1 mRNA isoforms in premutation carriers. J Med Genet 52(1):42-52.

Pretto, D. I., Kumar, M., Cao, Z., Cunningham, C. L., Durbin-Johnson, B., Qi, L., Berman, R., Noctor, S. C., Hagerman, R. J., Pessah, I. N., Tassone, F. (2014b). Reduced excitatory amino acid transporter 1 and metabotropic glutamate receptor 5 expression in the cerebellum of fragile X mental retardation gene 1 premutation carriers with fragile X-associated tremor/ataxia syndrome. Neurobiol Aging 35(5):1189-97.

Pretto, D. I., Mendoza-Morales, G., Lo, J., Cao, R., Hadd, A., Latham, G. J., Durbin-Johnson, B., Hagerman, R., Tassone, F. (2014c). CGG allele size somatic mosaicism and methylation in FMR1 premutation alleles. J Med Genet 51(5):309-18.

Primerano, B., Tassone, F., Hagerman, R. J., Hagerman, P. J., Amaldi, F., Bagni, C. (2002). Reduced FMR1 mRNA translation efficiency in Fragile X patients with premutations. RNA 8(12):1482-1488.

Qurashi, A., Li, W., Zhou, J. Y., Peng, J., Jin, P. (2011). Nuclear accumulation of stress response mRNAs contributes to the neurodegeneration caused by Fragile X premutation rCGG repeats. PLoS Genet 7(6):e1002102.

Ranum, L. P., Cooper, T. A. (2006). RNA-mediated neuromuscular disorders. Annu Rev Neurosci 29:259-77.

Rodriguez-Revenga, L., Madrigal, I., Pagonabarraga, J., Xuncla, M., Badenas, C., Kulisevsky, J., Gomez, B., Mila, M. (2009). Penetrance of FMR1 premutation associated pathologies in fragile X syndrome families. Eur J Hum Genet 17(10):1359-62.

Ross-Inta, C., Omanska-Klusek, A., Wong, S., Barrow, C., Garcia-Arocena, D., Iwahashi, C., Berry-Kravis, E., Hagerman, R. J., Hagerman, P. J., Giulivi, C. (2010). Evidence of mitochondrial dysfunction in fragile X-associated tremor/ataxia syndrome. Biochem J 429(3):545-52.

Rousseau, F., Rouillard, P., Morel, M. L., Khandjian, E. W., Morgan, K. (1995). Prevalence of carriers of premutation-size alleles of the FMRI gene--and implications for the population genetics of the fragile X syndrome. Am J Hum Genet 57(5):1006-18.

Santa Maria, L., Pugin, A., Alliende, M., Aliaga, S., Curotto, B., Aravena, T., Tang, H. T., Mendoza-Morales, G., Hagerman, R., Tassone, F. (2013). FXTAS in an unmethylated mosaic male with fragile X syndrome from Chile. Clin Genet.

Sellier, C., Freyermuth, F., Tabet, R., Tran, T., He, F., Ruffenach, F., Alunni, V., Moine, H., Thibault, C., Page, A., Tassone, F., Willemsen, R., Disney, M. D., Hagerman, P. J., Todd, P. K., Charlet-Berguerand, N. (2013). Sequestration of DROSHA and DGCR8 by expanded CGG RNA repeats alters microRNA processing in fragile X-associated tremor/ataxia syndrome. Cell Rep 3(3):869-80.

Sellier, C., Rau, F., Liu, Y., Tassone, F., Hukema, R. K., Gattoni, R., Schneider, A., Richard, S., Willemsen, R., Elliott, D. J., Adams, J. S. (2010). Sam68 sequestration and partial loss of function are associated with splicing alterations in FXTAS patients. EMBO J.

Sellier, C., Usdin, K., Pastori, C., Peschansky, V. J., Tassone, F., Charlet-Berguerand, N. (2014). The multiple molecular facets of fragile X-associated tremor/ataxia syndrome. J Neurodev Disord 6(1):23.

Sittler, A., Devys, D., Weber, C., Mandel, J. L. (1996). Alternative splicing of exon 14 determines nuclear or cytoplasmic localisation of *fmr1* protein isoforms. Human Molecular Genetics 5(1):95-102.

Sofola, O. A., Jin, P., Qin, Y., Duan, R., Liu, H., De Haro, M., Nelson, D. L., Botas, J. (2007). RNA-binding proteins hnRNP A2/B1 and CUGBP1 suppress fragile X CGG premutation repeat-induced neurodegeneration in a Drosophila model of FXTAS. Neuron 55(4):565-71.

Soontarapornchai, K., Maselli, R., Fenton-Farrell, G., Tassone, F., Hagerman, P. J., Hessl, D., Hagerman, R. J. (2008). Abnormal nerve conduction features in fragile X premutation carriers. Arch Neurol 65(4):495-8.

Tan, H., Poidevin, M., Li, H., Chen, D., Jin, P. (2012). MicroRNA-277 modulates the neurodegeneration caused by Fragile X premutation rCGG repeats. PLoS Genet 8(5):e1002681.

Tassone, F., Adams, J., Berry-Kravis, E. M., Cohen, S. S., Brusco, A., Leehey, M. A., Li, L., Hagerman, R. J., Hagerman, P. J. (2007a). CGG repeat length correlates with age of onset of motor signs of the fragile X-associated tremor/ataxia syndrome (FXTAS). Am J Med Genet B Neuropsychiatr Genet 144(4):566-9.

Tassone, F., Beilina, A., Carosi, C., Albertosi, S., Bagni, C., Li, L., Glover, K., Bentley, D., Hagerman, P. J. (2007b). Elevated FMR1 mRNA in premutation carriers is due to increased transcription. RNA 13(4):555-562.

Tassone, F., De Rubeis, S., Carosi, C., La Fata, G., Serpa, G., Raske, C., Willemsen, R., Hagerman, P. J., Bagni, C. (2011). Differential usage of transcriptional start sites and polyadenylation sites in FMR1 premutation alleles. Nucleic Acids Res 39(14):6172-85.

Tassone, F., Greco, C. M., Hunsaker, M. R., Seritan, A. L., Berman, R. F., Gane, L. W., Jacquemont, S., Basuta, K., Jin, L. W., Hagerman, P. J., Hagerman, R. J. (2012a). Neuropathological, clinical and molecular pathology in female fragile X premutation carriers with and without FXTAS. Genes Brain Behav 11(5):577-85.

Tassone, F., Hagerman, R. J., Garcia-Arocena, D., Khandjian, E. W., Greco, C. M., Hagerman, P. J. (2004). Intranuclear inclusions in neural cells with premutation alleles in fragile X associated tremor/ataxia syndrome. J Med Genet 41(4):e43.

Tassone, F., Hagerman, R. J., Loesch, D. Z., Lachiewicz, A., Taylor, A. K., Hagerman, P. J. (2000a). Fragile X males with unmethylated, full mutation trinucleotide repeat expansions have elevated levels of FMR1 messenger RNA. Am J Med Genet 94(3):232-6.

Tassone, F., Hagerman, R. J., Taylor, A. K., Gane, L. W., Godfrey, T. E., Hagerman, P. J. (2000b). Elevated levels of FMR1 mRNA in carrier males: a new mechanism of involvement in the fragile-X syndrome. Am J Hum Genet 66(1):6-15.

Tassone, F., Iong, K. P., Tong, T. H., Lo, J., Gane, L. W., Berry-Kravis, E., Nguyen, D., Mu, L. Y., Laffin, J., Bailey, D. B., Jr., Hagerman, R. J. (2012b). FMR1 CGG allele size and prevalence ascertained through newborn screening in the United States. Genome Med 4(12):100.

Tassone, F., Longshore, J., Zunich, J., Steinbach, P., Salat, U., Taylor, A. K. (1999). Tissue-specific methylation differences in a fragile X premutation carrier. Clinical Genetics 55(5):346-51.

Todd, P. K., Oh, S. Y., Krans, A., He, F., Sellier, C., Frazer, M., Renoux, A. J., Chen, K. C., Scaglione, K. M., Basrur, V., Elenitoba-Johnson, K., Vonsattel, J. P., Louis, E. D., Sutton, M. A., Taylor, J. P., Mills, R. E., Charlet-Berguerand, N., Paulson, H. L. (2013). CGG repeat-associated translation mediates neurodegeneration in fragile X tremor ataxia syndrome. Neuron 78(3):440-55.

Todd, P. K., Oh, S. Y., Krans, A., Pandey, U. B., Di Prospero, N. A., Min, K. T., Taylor, J. P., Paulson, H. L. (2010). Histone deacetylases suppress CGG repeat-induced neurodegeneration via transcriptional silencing in models of fragile X tremor ataxia syndrome. PLoS Genet 6(12):e1001240.

Toledano-Alhadef, H., Basel-Vanagaite, L., Magal, N., Davidov, B., Ehrlich, S., Drasinover, V., Taub, E., Halpern, G. J., Ginott, N., Shohat, M. (2001). Fragile-X Carrier Screening and the Prevalence of Premutation and Full-Mutation Carriers in Israel. Am J Hum Genet 69:351-360.

Tran, T., Childs-Disney, J. L., Liu, B., Guan, L., Rzuczek, S., Disney, M. D. (2014). Targeting the r(CGG) repeats that cause FXTAS with modularly assembled small molecules and oligonucleotides. ACS Chem Biol 9(4):904-12.

Tzeng, C. C., Tsai, L. P., Hwu, W. L., Lin, S. J., Chao, M. C., Jong, Y. J., Chu, S. Y., Chao, W. C., Lu, C. L. (2005). Prevalence of the FMR1 mutation in Taiwan assessed by large-scale screening of newborn boys and analysis of DXS548-FRAXAC1 haplotype. Am J Med Genet A 133(1):37-43.

Usdin, K., Hayward, B. E., Kumari, D., Lokanga, R. A., Sciascia, N., Zhao, X. N. (2014). Repeat-mediated genetic and epigenetic changes at the FMR1 locus in the Fragile X-related disorders. Front Genet 5:226.

Van Dam, D., Errijgers, V., Kooy, R. F., Willemsen, R., Mientjes, E., Oostra, B. A., De Deyn, P. P. (2005). Cognitive decline, neuromotor and behavioural disturbances in a mouse model for fragile-X-associated tremor/ataxia syndrome (FXTAS). Behav Brain Res 162(2):233-9.

Verheij, C., Bakker, C. E., De Graaff, E., Keulemans, J., Willemsen, R., Verkerk, A. J., Galjaard, H., Reuser, A. J., Hoogeveen, A. T., Oostra, B. A. (1993). Characterization and localization of the FMR-1 gene product associated with fragile X syndrome. Nature 363(6431):722-4.

Verheij, C., De Graaff, E., Bakker, C. E., Willemsen, R., Willems, P. J., Meijer, N., Galjaard, H., Reuser, A. J., Oostra, B. A., Hoogeveen, A. T. (1995). Characterization of FMR1 proteins isolated from different tissues. Hum Mol Genet 4(5):895-901.

Verkerk, A. J., De Graaff, E., De Boulle, K., Eichler, E. E., Konecki, D. S., Reyniers, E., Manca, A., Poustka, A., Willems, P. J., Nelson, D. L., Oostra, B.A. (1993). Alternative splicing in the fragile X gene FMR1. Hum Mol Genet 2(8):1348.

Verkerk, A. J., Pieretti, M., Sutcliffe, J. S., Fu, Y. H., Kuhl, D. P., Pizzuti, A., Reiner, O., Richards, S., Victoria, M. F., Zhang, F. P., Eussen, B. E., Vanommen, G. J. B., Blonden, L. a. J., Riggins, G. J., Chastain, J. L., Kunst, C. B., Galjaard, H., Caskey, T. C., Nelson, D. L., Oostra, B. A., Warren, S. T. (1991). Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. Cell 65(5):905-914.

Wang, Z., Taylor, A. K., Bridge, J. A. (1996). FMR1 fully expanded mutation with minimal methylation in a high functioning fragile X male. J Med Genet 33(5):376-8.

Wheeler, A. C., Mussey, J., Villagomez, A., Bishop, E., Raspa, M., Edwards, A., Bodfish, J., Bann, C., Bailey, D. B., Jr. (2015). DSM-5 changes and the prevalence of parent-reported autism spectrum symptoms in Fragile X syndrome. J Autism Dev Disord 45(3):816-29.

Willemsen, R., Hoogeveen-Westerveld, M., Reis, S., Holstege, J., Severijnen, L. A., Nieuwenhuizen, I. M., Schrier, M., Van Unen, L., Tassone, F., Hoogeveen, A. T., Hagerman, P. J., Mientjes, E. J., Oostra, B. A. (2003). The FMR1 CGG repeat mouse displays ubiquitin-positive intranuclear neuronal inclusions; implications for the cerebellar tremor/ataxia syndrome. Hum Mol Genet 12(9):949-59.

Yang, W. Y., Wilson, H. D., Velagapudi, S. P., Disney, M. D. (2015). Inhibition of Non-ATG Translational Events in Cells via Covalent Small Molecules Targeting RNA. J Am Chem Soc 137(16):5336-45.

Yao, B., Lin, L., Street, R. C., Zalewski, Z. A., Galloway, J. N., Wu, H., Nelson, D. L., Jin, P. (2014). Genome-wide alteration of 5-hydroxymethylcytosine in a mouse model of fragile X-associated tremor/ataxia syndrome. Hum Mol Genet 23(4):1095-107.

Yrigollen, C. M., Durbin-Johnson, B., Gane, L., Nelson, D. L., Hagerman, R., Hagerman, P. J., Tassone, F. (2012). AGG interruptions within the maternal FMR1 gene reduce the risk of offspring with fragile X syndrome. Genet Med 14(8):729-36.

Yrigollen, C. M., Martorell, L., Durbin-Johnson, B., Naudo, M., Genoves, J., Murgia, A., Polli, R., Zhou, L., Barbouth, D., Rupchock, A., Finucane, B., Latham, G. J., Hadd, A., Berry-Kravis, E., Tassone, F. (2014). AGG interruptions and maternal age affect FMR1 CGG repeat allele stability during transmission. J Neurodev Disord 6(1):24.

Yu, S., Pritchard, M., Kremer, E., Lynch, M., Nancarrow, J., Baker, E., Holman, K., Mulley, J. C., Warren, S. T., Schlessinger, D., Et Al. (1991). Fragile X genotype characterized by an unstable region of DNA. Science 252(5010):1179-81.

Zhang, K., Donnelly, C. J., Haeusler, A. R., Grima, J. C., Machamer, J. B., Steinwald, P., Daley, E. L., Miller, S. J., Cunningham, K. M., Vidensky, S., Gupta, S., Thomas, M. A., Hong, I., Chiu, S. L., Huganir, R. L., Ostrow, L. W., Matunis, M. J., Wang, J., Sattler, R., Lloyd, T. E., Rothstein, J. D. (2015). The C9orf72 repeat expansion disrupts nucleocytoplasmic transport. Nature 525(7567):56-61.

Zhong, N., Ju, W., Pietrofesa, J., Wang, D., Dobkin, C., Brown, W. T. (1996). Fragile X "gray zone" alleles: AGG patterns, expansion risks, and associated haplotypes. Am J Med Genet 64(2):261-5.

Zhong, N., Yang, W., Dobkin, C., Brown, W. T. (1995). Fragile X gene instability: anchoring AGGs and linked microsatellites [see comments]. Am J Hum Genet 57(2):351-61.

Zu, T., Gibbens, B., Doty, N. S., Gomes-Pereira, M., Huguet, A., Stone, M. D., Margolis, J., Peterson, M., Markowski, T. W., Ingram, M. A., Nan, Z., Forster, C., Low, W. C., Schoser, B., Somia, N. V., Clark, H. B., Schmechel, S., Bitterman, P. B., Gourdon, G., Swanson, M. S., Moseley, M., Ranum, L. P. (2011). Non-ATG-initiated translation directed by microsatellite expansions. Proc Natl Acad Sci U S A 108(1):260-5.

Zu, T., Liu, Y., Banez-Coronel, M., Reid, T., Pletnikova, O., Lewis, J., Miller, T. M., Harms, M. B., Falchook, A. E., Subramony, S. H., Ostrow, L. W., Rothstein, J. D., Troncoso, J. C., Ranum, L. P. (2013). RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. Proc Natl Acad Sci U S A 110(51):E4968-77.

Zuhlke, C., Budnik, A., Gehlken, U., Dalski, A., Purmann, S., Naumann, M., Schmidt, M., Burk, K., Schwinger, E. (2004). FMR1 premutation as a rare cause of late onset ataxia--evidence for FXTAS in female carriers. J Neurol 251(11):1418-1419.

The EMBO Journal (2010) 29, 1248–1261 | © 2010 European Molecular Biology Organization | All Rights Reserved 0261-4189/10
www.embojournal.org

THE EMBO JOURNAL

Sam68 sequestration and partial loss of function are associated with splicing alterations in FXTAS patients

Chantal Sellier[1], Frédérique Rau[1], Yilei Liu[2], Flora Tassone[3,6], Renate K Hukema[4], Renata Gattoni[1], Anne Schneider[1], Stéphane Richard[5], Rob Willemsen[4], David J Elliott[2], Paul J Hagerman[3,6] and Nicolas Charlet-Berguerand[1,*]

[1]Department of Neurobiology and Genetics, IGBMC, INSERM U964, CNRS UMR7104, University of Strasbourg, Illkirch, France, [2]Institute of Human Genetics, Newcastle University, Newcastle, UK, [3]M.I.N.D. Institute, University of California, Davis, Health System, Sacramento, CA, USA, [4]CBG-Department of Clinical Genetics, Erasmus MC, Rotterdam, The Netherlands, [5]Lady Davis Institute for Medical Research, Departments of Medicine and Oncology, McGill University, Montreal, Quebec, Canada and [6]Department of Biochemistry and Molecular Medicine, School of Medicine, University of California, Davis, CA, USA

Fragile X-associated Tremor/Ataxia Syndrome (FXTAS) is a neurodegenerative disorder caused by expansion of 55–200 CGG repeats in the 5′-UTR of the *FMR1* gene. FXTAS is characterized by action tremor, gait ataxia and impaired executive cognitive functioning. It has been proposed that FXTAS is caused by titration of RNA-binding proteins by the expanded CGG repeats. Sam68 is an RNA-binding protein involved in alternative splicing regulation and its ablation in mouse leads to motor coordination defects. Here, we report that mRNAs containing expanded CGG repeats form large and dynamic intranuclear RNA aggregates that recruit several RNA-binding proteins sequentially, first Sam68, then hnRNP-G and MBNL1. Importantly, Sam68 is sequestered by expanded CGG repeats and thereby loses its splicing-regulatory function. Consequently, Sam68-responsive splicing is altered in FXTAS patients. Finally, we found that regulation of Sam68 tyrosine phosphorylation modulates its localization within CGG aggregates and that tautomycin prevents both Sam68 and CGG RNA aggregate formation. Overall, these data support an RNA gain-of-function mechanism for FXTAS neuropathology, and suggest possible target routes for treatment options.

The EMBO Journal (2010) **29,** 1248–1261. doi:10.1038/emboj.2010.21; Published online 25 February 2010
Subject Categories: RNA; molecular biology of disease
Keywords: FXTAS; RNA gain-of-function diseases; Sam68

*Corresponding author. Department of Neurobiology and Genetics, IGBMC, 1 rue Laurent Fries, Strasbourg, Illkirch 67400, France. Tel.: +33 388 653 309, Fax: +33 388 653 201; E-mail: ncharlet@igbmc.fr

Received: 29 May 2009; accepted: 20 January 2010; published online: 25 February 2010

Introduction

Fragile X-associated Tremor/Ataxia Syndrome (FXTAS) is a recently identified neurodegenerative disorder that affects principally older adult males who are carriers of pre-mutation expansions (55–200 CGG repeats) in the 5′-untranslated region (UTR) of the Fragile X Mental Retardation-1 (*FMR1*) gene (Hagerman *et al*, 2001; Hagerman and Hagerman, 2004; Ostra and Willemesen, 2009). The major clinical features of FXTAS are progressive intention tremor and gait ataxia, with more variable associated features, including parkinsonism, dysautonomia, anxiety, peripheral neuropathy and cognitive decline (Jacquemont *et al*, 2003). The neuropathological hallmark of FXTAS is the presence of ubiquitin-positive intranuclear inclusions in both astrocytes and neurons throughout the brain (Greco *et al*, 2002). It is estimated that nearly 1 in 3000 males has a lifetime risk of developing FXTAS, which would make FXTAS one of the most common single gene causes of tremor, ataxia and cognitive decline among older adults (Jacquemont *et al*, 2004).

In Fragile X syndrome full mutations (>200 CGG repeats) result in hypermethylation and silencing of the *FMR1* gene. In contrast, carriers of the pre-mutation alleles (55–200 CGG repeats) have increased *FMR1* mRNA levels but normal, or moderately low, FMR1 protein expression, especially in the upper pre-mutation range (Tassone *et al*, 2000a, b; Kenneson *et al*, 2001; Primerano *et al*, 2002). These observations, coupled with the fact that mRNAs containing expanded CGG repeats accumulate in intranuclear aggregates, suggest a toxic RNA gain-of-function model for FXTAS (Tassone *et al*, 2004). In support of this hypothesis, a knock-in (KI) mouse model, in which the endogenous eight CGG repeats in the *Fmr1* gene has been replaced with an expansion containing ~100 CGG repeats of human origin, shows ubiquitin-positive intranuclear inclusions and mild neuromotor and behavioural disturbance (Willemsen *et al*, 2003; Van Dam *et al*, 2005; Brouwer *et al*, 2008). Furthermore, sole expression of mRNAs containing 90 CGG repeats outside the context of *Fmr1* in a transgenic mouse model is sufficient to recapitulate the neuropathological and molecular features of FXTAS (Hashem *et al*, 2009). Similarly, heterologous expression of 90 CGG repeats in *Drosophila* is sufficient to cause neurodegeneration along with formation of neuronal inclusions (Jin *et al*, 2003). These models show that sole expression of expanded CGG repeats is necessary and sufficient to cause a pathology similar to human FXTAS, and thus indicate that the expanded CGG repeats in RNA are the likely cause of the neurodegeneration in FXTAS.

The FXTAS toxic RNA gain-of-function model show similarities with Myotonic Dystrophies (DM), where expanded CUG or CCUG repeats accumulate in nuclear RNA aggregates that sequester the Muscleblind-like (MBNL1) splicing factor. In DM, partial depletion of the free pool of MBNL1 leads to

specific alternative splicing changes, which ultimately result in the symptoms of DM (Ranum and Cooper, 2006; Wheeler and Thornton, 2007). Extending this model to FXTAS, expanded CGG repeats are predicted to sequester specific proteins resulting in loss of their normal molecular functions. Several proteins, including a number of heat-shock proteins, Purα, hnRNP-A2/B1, CUGBP1, MBNL1, lamin-A/C and MBP were found to localize with ubiquitin-positive inclusions in CGG-expressing *Drosophila*, KI mouse model and FXTAS patients (Iwahashi *et al*, 2006; Jin *et al*, 2007; Sofola *et al*, 2007). However, these proteins were not found to be sequestered by expanded CGG repeats and consequently they are not expected to lose their functions in FXTAS patients.

In this study, we found that in contrast to CUG repeats, expanded CGG repeats accumulate in dynamic intranuclear RNA structures that expand over time. Formation of giant CGG RNA aggregates ultimately results in disorganization of the nuclear lamin structure and cell death. MBNL1 and hnRNP-G proteins were found localized within CGG aggregates but only in the larger inclusions and at late time points after transfection, suggesting these are not the primary defects. In contrast, we identified the Src-Associated substrate during mitosis of 68-kDa (Sam68) protein as the only protein that colocalizes with CGG RNA aggregates at each time point. Sam68 is a nuclear RNA-binding protein involved in alternative splicing regulation (Stoss *et al*, 2001; Paronetto *et al*, 2007; Chawla *et al*, 2009), and its ablation in a mouse knock-out model leads to motor coordination defects (Lukong and Richard, 2008). Sam68 splicing activity, RNA-binding ability and localization are regulated by phosphorylation (Haegebarth *et al*, 2004; Lukong *et al*, 2005), and Sam68 interacts with various RNA-binding proteins through several protein–protein interaction domains (Lukong and Richard, 2003). We found that Sam68 is required for subsequent aggregation of MBNL1 and hnRNP G proteins within CGG aggregates. Most importantly, Sam68 is sequestered by expanded CGG-repeat aggregates and thereby loses its splicing-regulatory function. As a consequence, Sam68-regulated splicing is altered in FXTAS patients. Finally, we found that regulation of Sam68 phosphorylation modulates its localization within CGG aggregates. Strikingly, among the various kinase and phosphatase inhibitors tested, we found one, tautomycin, which not only prevents Sam68 colocalization within CGG aggregates, but also abolishes CGG RNA aggregate formation.

Results

Expanded CGG repeats form dynamic nuclear RNA aggregates that expand over time

We first questioned whether pre-mutation-length CGG repeats can form nuclear RNA aggregates in cultured cells. We transfected plasmids expressing either 20, 40, 60 or 100 CGG repeats under the control of a cytomegalovirus (CMV) promoter in various cell lines, and tested the formation of CGG aggregates by RNA fluorescence *in situ* hybridation (FISH) analysis. We confirmed the specificity of the FISH conditions, and the RNA composition of CGG aggregates, as they were sensitive to RNAse treatment (Supplementary Figure S1). Consistent with an RNA gain-of-function model, expression of 60 or 100 CGG repeats within COS7 cells generated numerous intranuclear CGG aggregates, whereas expression of 20 CGG repeats did not (Figure 1A). Expression of 40 CGG repeats resulted in an intermediate condition with formation of rare small intranuclear aggregates. This is consistent with observations in FXTAS patients in whom it is estimated that 'normal' CGG polymorphic repeat lengths are 5–45 repeats long, 'gray zone' alleles contain 45 to 55 repeats and FXTAS patients are defined by pre-mutation allele containing 55–200 CGG repeats (Tassone *et al*, 2007; Leehey *et al*, 2008).

Next, expression of CGG aggregates was investigated in various cell types. Transfection of constructs containing 60 CGG repeats led to intranuclear CGG RNA aggregate formation in primary cultures of hippocampal embryonic mouse neurons, as well as in various immortalized cell lines such as neuronal (differentiated PC12), ovarian (SKOV3 and SW626) and kidney (COS7)-derived cell lines (Figure 1B). However, no aggregates were observed in A172, U-937, THP1, HeLa, HEK293, NG108-15, IMR-32, Neuro-2a, SH-SY5Y, SK-N-MC or SK-N-SH cell lines (data not shown), confirming a previous report that not all cell lines can support CGG-repeat aggregate formation (Arocena *et al*, 2005). Tests of colocalization with various nuclear structures indicated that in transfected cells most of the CGG aggregates are associated with nuclear speckles, but not with other structures such as lamin, nucleoli, PML, Gems or Cajal bodies (Supplementary Figure S2A).

Finally, kinetic formation of CGG aggregates was investigated. COS7 cells were transfected with a plasmid expressing 60 CGG repeats and analysed by RNA FISH either 24, 48 or 72 h after transfection. Surprisingly, expanded CGG repeats formed dynamic nuclear structures that expanded over time, resulting in giant inclusions, nuclear lamin architecture disruption and cell death 72–96 h after transfection (Figure 1C). In contrast, expanded CUG (DM1 mutation) or AUUCU (SCA10 mutation) repeats were less toxic and formed stable nuclear aggregates, whose size did not evolve over time (Supplementary Figure S2B). Annexin-V–PE apoptosis tests were negative indicating that the cytotoxicity of CGG repeats is not linked to apoptotic cell death (data not shown), and in agreement with a previous report (Arocena *et al*, 2005) no ubiquitin-positive aggregates were observed in transfected cells.

CGG aggregates recruit Sam68, then MBNL1 and hnRNP G

To identify which proteins are associated with expanded CGG repeats, we first adopted an *in vitro* approach. Proteins extracted from mouse brain or COS7-cell nuclei were captured on streptavidin resin coupled to biotinylated *in vitro*-transcribed RNA composed of 60 CGG repeats, eluted, separated on SDS–PAGE gels and identified by MALDI-TOF analysis. More than 20 proteins were identified (Supplementary Table 1), including a heat-shock protein and several RNA-binding proteins, including MBNL1 and hnRNP-G.

To discard non-specific binding proteins, we tested for colocalization of these candidates with RNA aggregates in COS7 cells transfected with 60 CGG repeats. CGG aggregates expand over time, suggesting that these repeats may recruit different proteins at different time points. Thus, we tested our candidates at 24 and 72 h after transfection (Figure 2A and B). Colocalization of MBNL1 within CGG aggregates increased from 14% at 24 h to 41% at 72 h after transfection. Similarly,

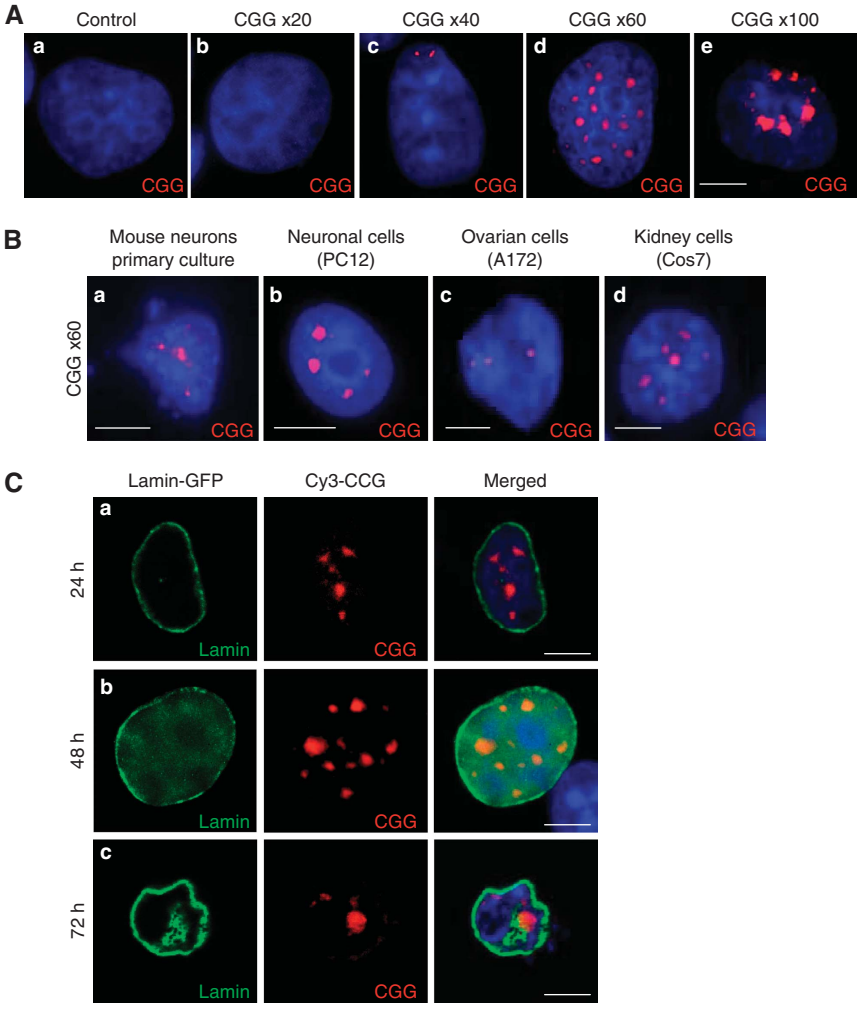


Figure 1 Expanded CGG repeats form intranuclear RNA aggregates. (**A**) COS7 cells were transfected with a plasmid expressing either no (a), 20 (b), 40 (c), 60 (d) or 100 (e) CGG repeats, transferred to 0.1% serum to block cell divisions and analysed 24 h after transfection by RNA FISH using a $(CCG)_{8x}$-Cy3 DNA probe. (**B**) Primary cultures of hippocampal embryonic mouse neurons (a), differentiated PC12 (b), SKOV3 (c) and COS7 (d) cells were transfected with a plasmid expressing 60 CGG repeats and analysed 24 h after transfection by FISH. (**C**) COS7 cells were co-transfected with a plasmid expressing 60 CGG repeats and a plasmid expressing GFP-tagged lamin-A, and analysed by RNA FISH either 24 (a), 48 (b) or 72 (c) hours after transfection. In all the figures, the magnification is $\times 630$. The scale bars represent 10 μm; nuclei were counterstained with DAPI and one representative experiment from at least three separate experiments is shown.

colocalization of hnRNP-G increased from 26 to 73% (Figure 2D), suggesting that CGG aggregates form dynamic structures, which constantly recruit proteins.

By contrast, other *in vitro*-identified candidates such as SPNR, hnRNP-A1, hnRNP-A2/B, hnRNP-C, hnRNP-D, hnRNP-E and hnRNP-H showed no or very little colocalization with RNA aggregates, and any colocalization observed was only within the giant CGG aggregates that form just before cell death (data not shown). We also observed that neither CUGBP1, nor Purα, colocalized with CGG aggregates in COS7-transfected cells.

These results suggest that MBNL1 and hnRNP-G are not initially recruited by CGG repeats, but join the larger aggregates later on, probably through protein–protein interactions. On the basis of that hypothesis, we searched for proteins that would colocalize with expanded CGG repeats early after transfection, and may capture other RNA-binding proteins through RNA or protein interactions. We screened ~50 candidates known to bind to RNA or RNA-binding proteins (Supplementary Table 2 and data not shown) and found one, Sam68, that consistently co-localized with CGG RNA aggregates at each time point, including the earliest (Figure 2C and D).

Next, we confirmed by FISH/IF that endogenous Sam68 colocalized with CGG aggregates in neuronal-differentiated PC12 cells (Figure 2C). Similar to COS7 cells, endogenous MBNL1 and hnRNP-G were not recruited within CGG aggregates in PC12 cells 24 h after transfection. We were not able to test Sam68, MBNL1 or hnRNP-G colocalization within CGG aggregates at later transfection time points as PC12 cells are very sensitive to CGG toxicity and die after less than 48 h of CGG expression.

Sam68 is a nuclear protein involved in various aspects of mRNA metabolism and interacts with several RNA-binding proteins, raising question its specificity towards CGG aggregates versus other expanded RNA repeats. COS7 cells were transfected by plasmids expressing either expanded CGG, CUG, CCUG or AUUCU repeats, which are the cause of FXTAS, DM1, DM2 and SCA10 diseases, respectively. We found that 24 or 72 h after transfection, endogenous Sam68

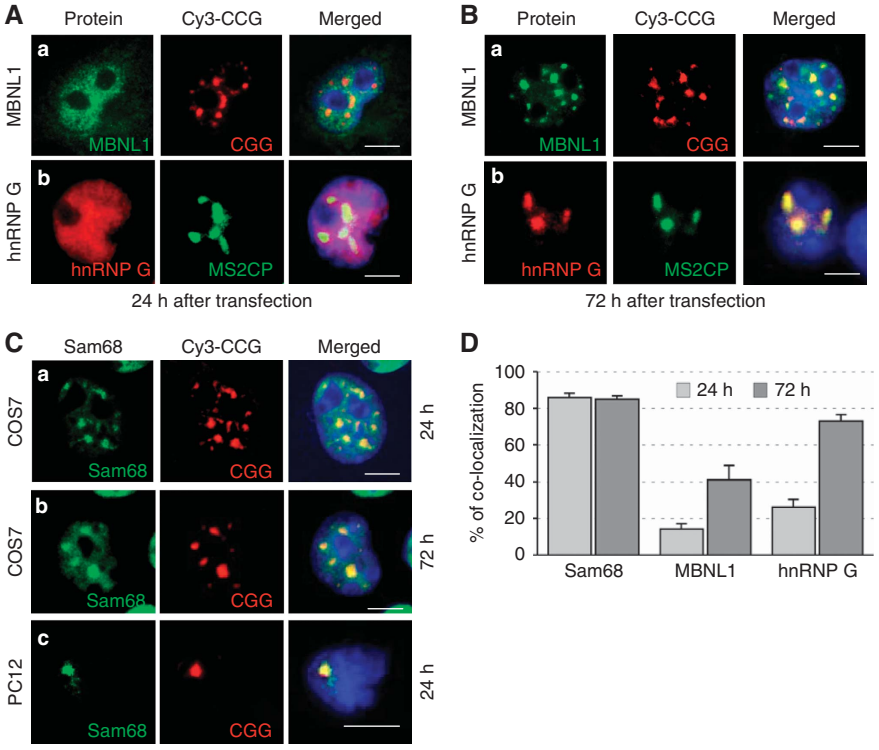


Figure 2 Sam68 colocalizes with CGG RNA aggregates. (**A**) COS7 cells were transfected with a plasmid expressing 60 CGG repeats and analysed by FISH/IF using an antibody against MBNL1 (a) 24 h after transfection. None of the antibodies against hnRNP-G that we tested supported FISH conditions. Consequently, endogenous hnRNP-G (b) was analysed by co-transfection of COS7 cells with a plasmid expressing 60 CGG repeats fused to three MS2 tags and a plasmid expressing the GFP-MS2 Coat Protein. The MS2 Coat Protein (MS2CP) possesses a very high and specific affinity for MS2 RNA tags. Endogenous hnRNP-G (b) was detected by IF and the CGG aggregates by localization of the GFP-MS2CP protein, which is bound to the MS2-$(CGG)_{60x}$ RNA. In the absence of MS2-$(CGG)_{60x}$, GFP-MS2CP was diffuse in the nucleoplasm (data not shown). (**B**) Similar to panel A but analysed 72 h after transfection. (**C**) COS7 cells (a, b) or differentiated PC12 neuronal cells (c) were transfected with a plasmid expressing 60 CGG repeats and analysed by FISH/IF using an antibody against Sam68, 24 h (a and c) or 72 h after transfection (b). (**D**) The percentage of endogenous Sam68, MBNL1 and hnRNP-G colocalized within CGG RNA aggregates in transfected COS7 cells 24 or 72 h after transfection. In all the experiments, three independent transfections totalling a hundred cells were counted, and results are presented as mean ± s.d.

colocalized only with CGG RNA aggregates, but not with other expanded RNA repeats (Figure 3). These results suggest that Sam68 is a specific and early component of CGG aggregates in FXTAS.

Sam68 is essential for recruitment of MBNL1 and hnRNP-G within CGG aggregates

Sam68 is recruited earlier than MBNL1 and hnRNP-G, suggesting sequential recruitment of proteins within the CGG aggregates. Furthermore, Sam68 protein contains several protein–protein interaction domains, raising the question whether Sam68 directly recruits hnRNP-G and MBNL1. In agreement with a previous report (Venables *et al*, 1999), we found by co-immunoprecipitation that Sam68 interacts with hnRNP-G. However, we found no robust interactions between MBNL1 and Sam68 or between MBNL1 and hnRNP-G (Supplementary Figure S3A).

Next, we mapped the domain required for Sam68 colocalization with CGG aggregates. Sam68 protein contains a central KH RNA-binding domain, and N- and C-terminal protein–protein interaction domains. Deletion of the Sam68 N-terminal domain abolished colocalization with CGG aggregates, whereas deletion of the Sam68 RNA-binding domain did not (Supplementary Figure S3B). Similarly, Sam68 paralog proteins, Slm1 and Slm2 (T-Star), which are devoid of the N-terminal extended region of the Sam68 protein, did not colocalize with expanded CGG aggregates 24 h after transfection. The N-terminal part of the Sam68 protein consists of a number of potential protein–protein interaction domains, suggesting that association of Sam68 with CGG repeats might not be direct but mediated through protein–protein interactions.

Finally, we tested whether Sam68 is required for colocalization of MBNL1 and hnRNP-G within CGG aggregates. Transfection of COS7 cells with an shRNA directed against Sam68 greatly reduced (>80%) the expression of endogenous Sam68 (Figure 4C). Importantly, depletion of endogenous Sam68 abolished the colocalization of MBNL1 and hnRNP-G within CGG aggregates (Figure 4 and Supplementary Table 3). These data show that Sam68 protein is required for subsequent recruitment of MBNL1 and hnRNP-G proteins within CGG aggregates.

Sam68 colocalizes with CGG aggregates in FXTAS patients

Next, we tested whether Sam68 protein colocalizes with endogenous CGG aggregates. We first analysed the localization of endogenous Sam68 and CGG repeats in the brain sections of a KI mouse model, in which endogenous CGG repeats had been replaced with an expansion of 98 CGG

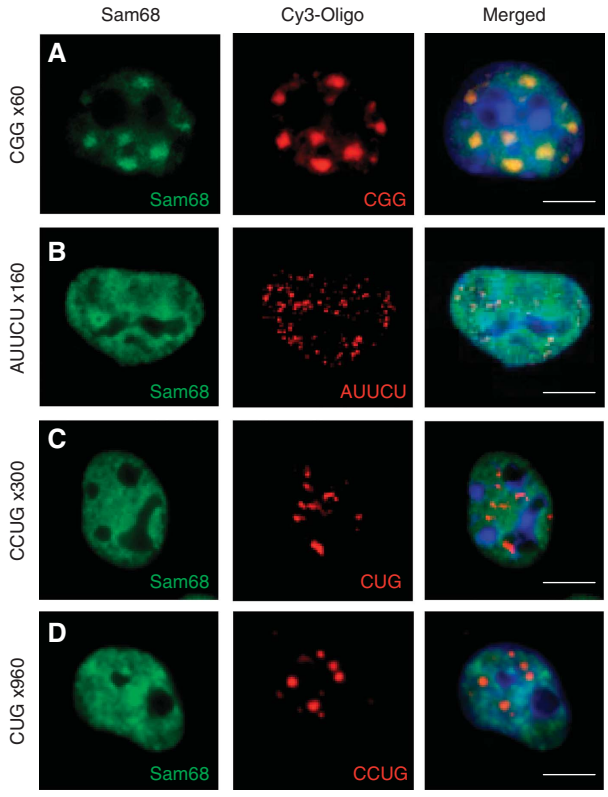


Figure 3 Sam68 colocalization within nuclear RNA aggregates is specific to expanded CGG repeats. COS7 cells were transfected with a plasmid expressing either 60 CGG repeats (**A**), 160 AUUCU repeats (**B**), 960 CUG repeats (**C**) or 300 CCUG repeats (**D**). Endogenous Sam68 and CGG repeats were analysed 24 h after transfection by FISH/IF.

repeats (Willemsen *et al*, 2003). FISH/IF experiments showed the presence of intranuclear CGG aggregates that colocalized with Sam68 in mice expressing 98 CGG repeats (Figure 5A). By contrast, no RNA aggregates were detected and Sam68 was diffuse throughout the nucleoplasm in control mice (Figure 5A).

Then, we investigated the presence of Sam68 and CGG aggregates in FXTAS patients. FISH/IF experiments show that Sam68 consistently colocalized with CGG intranuclear RNA aggregates in the brain sections of FXTAS patients (Figure 5B), but neither Sam68 nor CGG aggregates were found in control brain tissue (Figure 5B). We noted that in FXTAS patients, Sam68 and CGG aggregates were larger and more frequent than in KI mice. This is consistent with the milder neuromotor and behavioural disturbances observed in KI mice as compared with that in FXTAS patients (Willemsen *et al*, 2003; Van Dam *et al*, 2005).

Finally, presence of Sam68 aggregates in brain sections of FXTAS patients was confirmed by immunohistochemistry. In brain sections of control patients, Sam68 was diffuse within the nucleoplasm and no detectable Sam68 inclusions were found. In contrast, large intranuclear aggregates of Sam68 were observed in FXTAS patients (Figure 5C), thus confirming that CGG expanded repeats alter Sam68 localization.

Sam68 protein is partially immobilized within CGG expanded repeats

An RNA gain-of-function model for FXTAS predicts that CGG expanded repeats should immobilize Sam68 protein and deplete its molecular activity. To test Sam68 sequestration by CGG repeats, we first analysed its mobility by FRAP (Fluorescence Recovery After Photobleaching) experiments. FRAP of transfected GFP-Sam68 was measured in nuclear regions containing Sam68 aggregates and compared to nuclear regions containing diffuse Sam68 either located within the same nucleus or located in the nuclei of cells not transfected with CGG repeats (Figure 6A). In both cases, nucleoplasmic areas without Sam68 aggregates recovered ~95% of their initial fluorescence after photobleaching, whereas areas containing Sam68 aggregates recovered only ~60% of their initial fluorescence (Figure 6B). This shows that a fraction of Sam68 is less mobile in CGG-transfected cells.

Sam68 depletion by CGG repeats affects alternative splicing

Sam68 is a nuclear RNA-binding protein with roles in alternative splicing regulation (Stoss *et al*, 2001; Paronetto *et al*, 2007; Chawla *et al*, 2009). According to the RNA gain-of-function model, titration of free nuclear Sam68 into CGG nuclear aggregates should deplete its functional activity and result in detectable pre-mRNA splicing alterations. To test this hypothesis, we co-transfected constructs encoding expanded CGG repeats with minigenes that recapitulate splicing events directly regulated by Sam68.

In agreement with a previous report (Paronetto *et al*, 2007), overexpression of Sam68 with a *Bcl-x* minigene repressed the formation of the long splicing form, *Bcl-xL*. We found that overexpression of expanded CGG repeats reproduced a depletion of Sam68 and stimulated the expression of *Bcl-xL* (Figure 7A). Next, Sam68 paralogues SLM1 and SLM2 are known to regulate the alternative splicing of exon-7 of the survival motor neuron-2 (*SMN2*) pre-mRNA (Stoss *et al*, 2004); thus we tested whether Sam68 also regulates *SMN2* splicing. We found that overexpression of Sam68 modestly repressed the inclusion of the exon-7 of an SMN2 minigene, while expression of CGG repeats reproduced a depletion of Sam68 and stimulated the inclusion of that exon (Figure 7B). Finally, in a bioinformatic analysis, we identified a novel exon in intron-28 of the human *ATP11B* gene, which was predicted to be specifically included in the central neural system (Clark *et al* 2007; Liu *et al*, 2009). We constructed a minigene in which *ATP11B* exon-28B and ~300 bp of its flanking introns were bordered by β-globin exons and co-transfection experiments showed that Sam68 activated its inclusion. In contrast, overexpression of CGG repeats reproduced a depletion of endogenous Sam68 by shRNA and repressed exon-28B inclusion (Figure 7C).

hnRNP-G and MBNL1 are also recruited within CGG aggregates, raising the question whether the splicing defects observed in CGG expressing cells are only due to Sam68 sequestration. To test that hypothesis, we used a mutant of Sam68 deleted of its N-terminal domain (Sam68ΔNter), which was no longer recruited within CGG aggregates (Supplementary Figure S3B), but was still able to regulate alternative splicing (Supplementary Figure S4). Co-transfection of Sam68ΔNter rescued the splicing defects caused by expression of CGG repeats (Supplementary Figure S4), suggesting that the splicing alterations observed in CGG-expressing cells were mostly due to sequestration of Sam68.

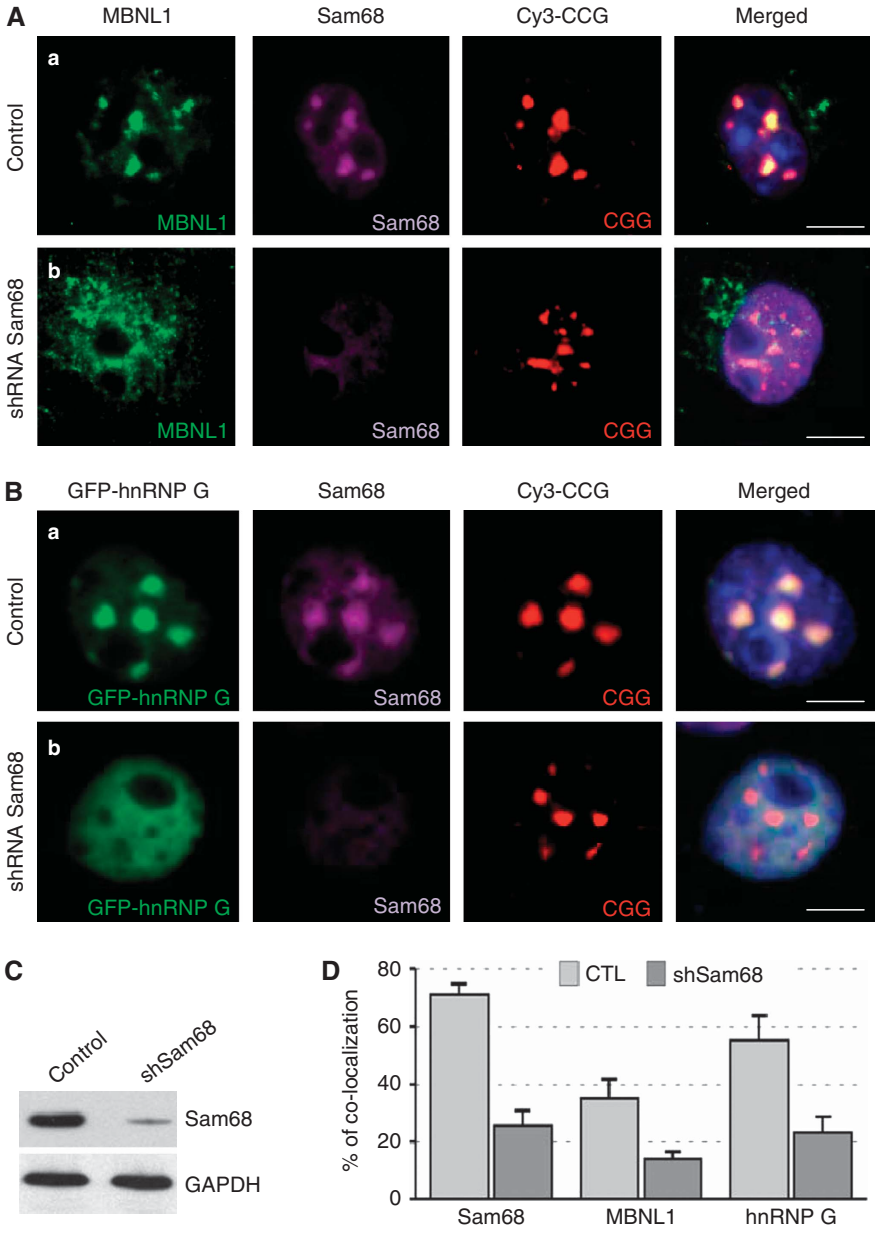


Figure 4 Sam68 is essential for recruitment of MBNL1 and hnRNP-G within CGG aggregates. (**A**) COS7 cells were co-transfected with a plasmid expressing 60 CGG repeats and a plasmid expressing either a control LacZ shRNA (a), or a Sam68 shRNA (b), and analysed 72 h after transfection by FISH/IF. Endogenous MBNL1 was detected by IF using an Alexa-488-labelled secondary antibody. Simultaneous detection of Sam68 by IF using a Cy5-labelled secondary antibody confirmed shRNA-mediated depletion of endogenous Sam68. (**B**) Endogenous hnRNP-G cannot be detected by FISH/IF. Thus, COS7 cells were co-transfected with a plasmid expressing 60 CGG repeats, a plasmid expressing GFP-hnRNP-G and either control (a) or Sam68 (b) shRNA and analysed 72 h after transfection by FISH. (**C**) shRNA-mediated depletion of endogenous Sam68 was confirmed by western blotting against Sam68. (**D**) The percentage of endogenous Sam68, MBNL1 and GFP-hnRNP-G colocalized within CGG RNA aggregates 72 h after transfection in COS7 cells transfected with a plasmid expressing either a control or a Sam68 shRNA.

Alternative splicing is altered in FXTAS patients

Sam68 has reduced nuclear spatial mobility and splicing regulatory function in CGG-expressing cells; so we tested whether alternative pre-mRNA splicing is altered in human FXTAS patients. RT–PCR analysis of *ATP11B* alternative splicing in brain samples of control and FXTAS patients showed a significant decrease of exon-28B inclusion from 47 ± 2% in control to 31 ± 8% in FXTAS patients ($P<0.005$). We confirmed these results by qRT–PCR and found that expression of *ATP11B* exon-28B is downregulated in FXTAS patients as compared with that in a control (Figure 8A). Similarly, we tested by qRT–PCR the expression of exon-7 of the *SMN2* pre-mRNA and found it under-expressed in FXTAS patients (Figure 8B). By contrast, we found no significant differences in the expression of *SMN1* exon-7, which is consistent with no alternative splicing regulation of that exon. These results are consistent with alternative splicing being altered in FXTAS patients.

Tyrosine phosphorylation of Sam68 regulates its recruitment within CGG aggregates

Sam68 is a substrate of the SIK/BRK-tyrosine kinase, which regulates the RNA-binding ability and localization of the Sam68 protein (Derry *et al*, 2000; Haegebarth *et al*, 2004;

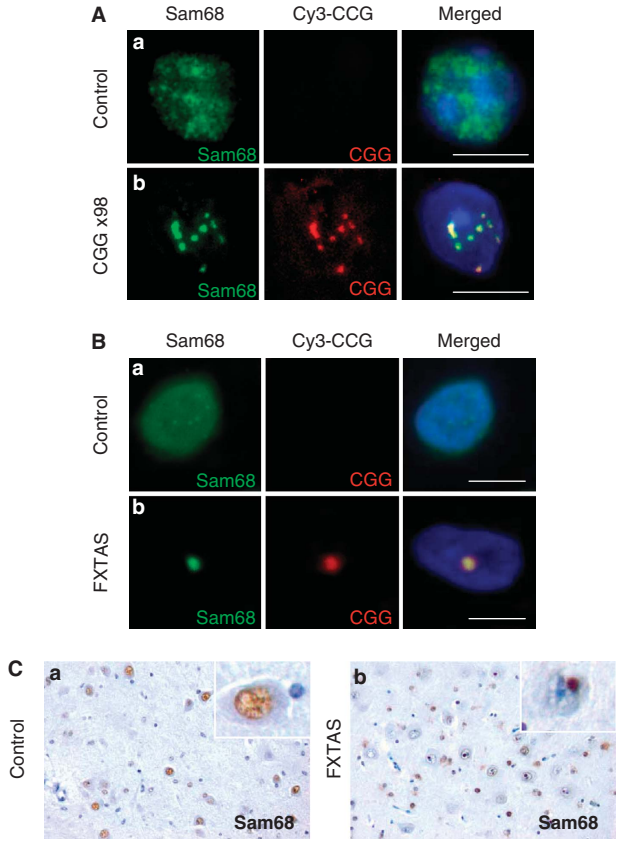


Figure 5 Sam68 colocalizes with endogenous CGG aggregates. (**A**) Brain sections of mouse expressing either control eight CGG repeats (a) or expanded 98 CGG repeats (b) were analysed by FISH/IF. (**B**) Similar to panel A. Brain sections (hippocampal area) of age-matched control (a) or FXTAS (b) patients were analysed by FISH/IF. Magnification: ×630. Scale bar, 10 μm. (**C**) Brain (hippocampal area) sections of age-matched control (a) or FXTAS (b) patients were analysed by immunohistochemistry directed against Sam68. Magnification: ×350.

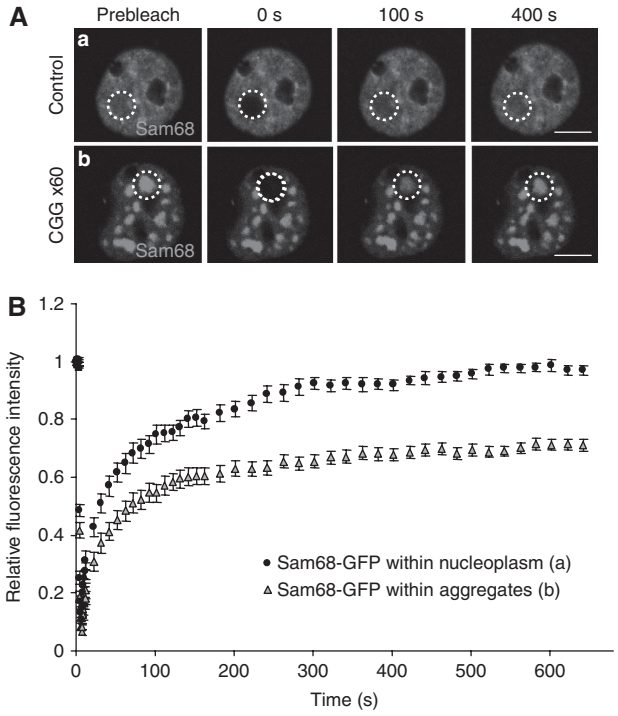


Figure 6 FRAP analysis uncovers an immobile fraction of GFP-Sam68 within CGG aggregates. (**A**) Photobleaching was performed 24 h after transfection of COS7 cells with GFP-Sam68 and a plasmid expressing either no CGG repeats (a) or 60 CGG repeats (b). The white circles denote the photobleached regions in the aggregates and the nucleoplasm. Representative images show a single *z*-section obtained before photobleaching (pre-bleach) and at the indicated time points after photobleaching. (**B**) Recovery curves of photobleached aggregates and nucleoplasm in cells expressing 0 or 60 CGG repeats are shown as relative fluorescence intensity versus time. In CGG-expressing cells, recovery reached a plateau at ~60% around 300 s. Each data point is the average of 10 nuclei. The error bars indicate the s.e.m.'s.

Lukong *et al*, 2005). We thus tested the effect of Sam68 phosphorylation on the formation of CGG aggregate. We found that expression of SIK/BRK disrupted Sam68 protein localization within CGG aggregates and returned Sam68 to a free nucleoplasmic localization (Figure 9A and B). Sam68 is tyrosine-phosphorylated in response to EGF treatment (Lukong *et al*, 2005). We found that inhibition of the EGFR tyrosine kinase pathways through tyrphostin/AG490 treatment stimulated the recruitment of Sam68 within CGG aggregates, leading to formation of large aggregates at early time points. By contrast, treatment of transfected cells with dephostatin, a tyrosine phosphatase inhibitor, reduced the recruitment of Sam68 within CGG aggregates (Figure 9A and B). These data suggest a model in which tyrosine phosphorylation of Sam68, mediated by the EGFR-SIK/BRK tyrosine kinase pathway, reduces the recruitment of Sam68 within CGG aggregates (Supplementary Figure S6A).

We confirmed that Sam68 phosphorylation was stimulated by transfection of the SIK/BRK kinase or by inhibition of tyrosine phosphatase pathways through dephostatin treatment. By contrast, inhibition of the EGFR pathway (tyrphostin/AG490 treatment) reduced Sam68 phosphorylation (Figure 9C).

Tyrosine phosphorylation of Sam68 inhibits its splicing activity (Paronetto *et al*, 2007). We found that transfection of SIK/BRK or dephostatin treatment, both of which induced Sam68 tyrosine phosphorylation (Figure 9C), modified the splicing of *ATP11B* and mimicked a depletion of Sam68 by CGG or shRNA expression (Figure 9D and data not shown). Interestingly, we noted no cumulative effects on splicing when CGG-expressing cells were co-transfected with SIK/BRK or treated with dephostatin. These data suggest that either *ATP11B* splicing regulation was saturated or that CGG repeats acted similarly to SIK/BRK or dephostatin, probably through inhibition of Sam68 activity. We observed a similar absence of cumulative effects when CGG repeats and Sam68 shRNA were coexpressed (data not shown), suggesting that CGG repeats acted on splicing mostly through Sam68 depletion.

Tautomycin reduces the formation of CGG aggregates

Among the various inhibitors of kinases and phosphatases that we tested, we found one, tautomycin, which not only reduced Sam68 colocalization but also reduced CGG aggregate formation (Figure 10A and Supplementary Table 3). By contrast, tautomycin treatment did not impair the formation

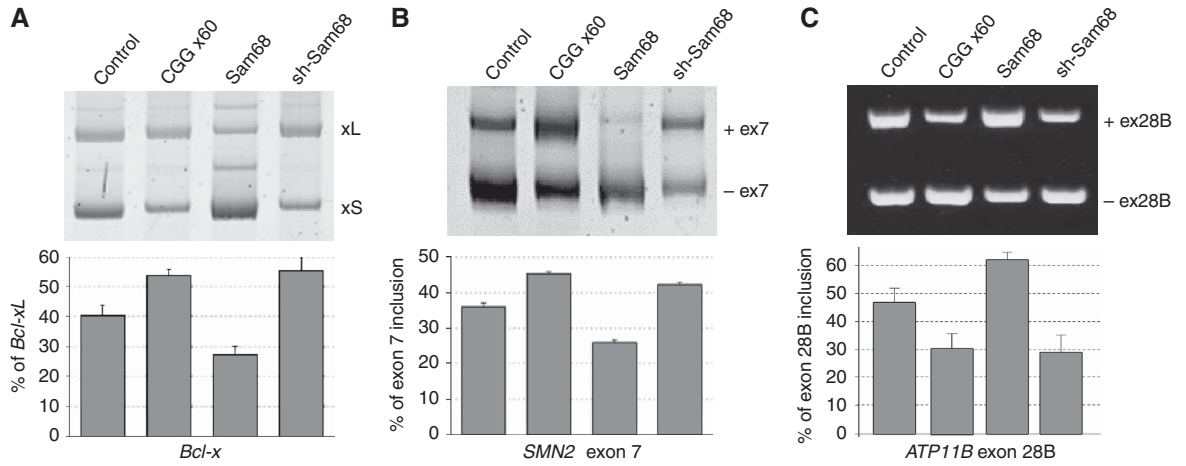


Figure 7 Alternative splicing is altered in CGG-expressing cells. COS7 cells were co-transfected with a *Bcl-x* minigene (**A**), an *SMN2* exon-7 minigene (**B**), or an *ATP11B* exon-28B minigene (**C**) and a plasmid expressing either no CGG repeat, or 60 CGG repeats, Sam68 or Sam68 shRNA. *Bcl-xL/S*, SMN2 and ATP11B splicing isoforms were identified 24 h after transfection by RNA extraction followed by RT–PCR.

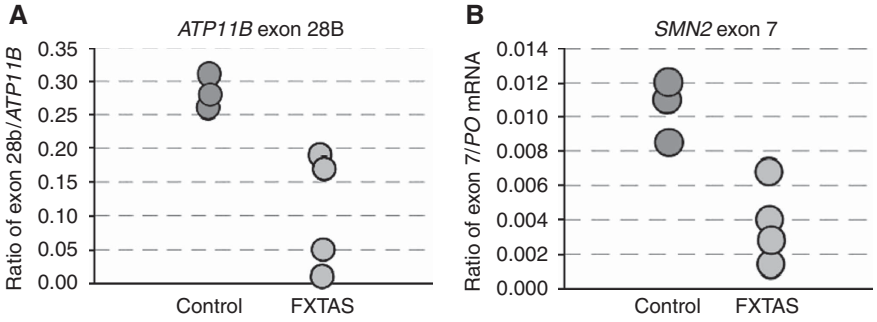


Figure 8 Alternative splicing is altered in FXTAS patients. (**A**) qRT–PCR analysis of *ATP11B* exon-28B in brain RNA samples of three controls and four FXTAS patients. Results are expressed as the ratio between alternative *ATP11B* exon-28B and constitutive *ATP11B* exons 21–22. (**B**) qRT–PCR analysis of *SMN2* exon-7 in brain RNA samples from three controls and four FXTAS patients. We were not able to discriminate between *SMN1* and *SMN2* constitutive exons; therefore, results are expressed as the ratio of the alternative *SMN2* exon-7/*PO* mRNA.

of expanded CUG- or AUUCU-repeat RNA aggregates (data not shown).

We confirmed by qRT–PCR that the quantity of RNA containing the expanded CGG repeats was not altered by tautomycin treatment (Figure 10C), suggesting that tautomycin disrupts CGG aggregation without altering their expression.

Finally, we observed that in the presence of tautomycin, *ATP11B* splicing was no longer altered by expression of expanded CGG repeats (Figure 10D), suggesting that tautomycin reduced the deleterious effects of the CGG repeats on splicing. However and as noted previously (Mermoud *et al*, 1992; Novoyatleva *et al*, 2008), treatment with tautomycin alone resulted in splicing changes (Figure 10D), rendering difficult to pinpoint the precise role of tautomycin on the deleterious effects of CGG repeats on splicing.

Discussion

CGG aggregates are dynamic structures that recruit various RNA-binding proteins sequentially

A striking characteristic of CGG expanded repeats is that they form dynamic intranuclear RNA aggregates that enlarge with time, resulting in the formation of giant inclusions, an observation that stands in contrast to the absence of growth over time of expanded CUG (DM1 mutation), CCUG (DM2) or AUUCU (SCA10) repeats. Continuous enlargement of CGG RNA aggregates suggests that these repeats may constantly recruit proteins.

Consistent with that hypothesis, we and others have found various proteins, mainly RNA-binding proteins, to be captured by CGG repeats *in vitro* (Iwahashi *et al*, 2006; Jin *et al*, 2007; Sofola *et al*, 2007), suggesting that CGG repeats can sequester a large number of proteins through direct RNA–protein interactions, but probably also through indirect protein–protein interactions. When tested in cells, we found that endogenous Sam68, MBNL1 and hnRNP-G colocalized with CGG aggregates. In contrast, other candidate proteins such as Purα, FMRP and CUGBP1 did not colocalize with expanded CGG RNA, whereas hnRNP-A1, hnRNP-A2/B, hnRNP-C, hnRNP-D, hnRNP-E and hnRNP-H presented an intermediary situation with some colocalization, but only at very late time points and within the giant CGG aggregates that form in dying cells. This questions whether these other candidate proteins are recruited specifically at the very last step of CGG aggregate formation, or that they form non-specific aggregates in dying cells. We tested the localization of some of these candidates by FISH/IF in brain sections of control and FXTAS patients.

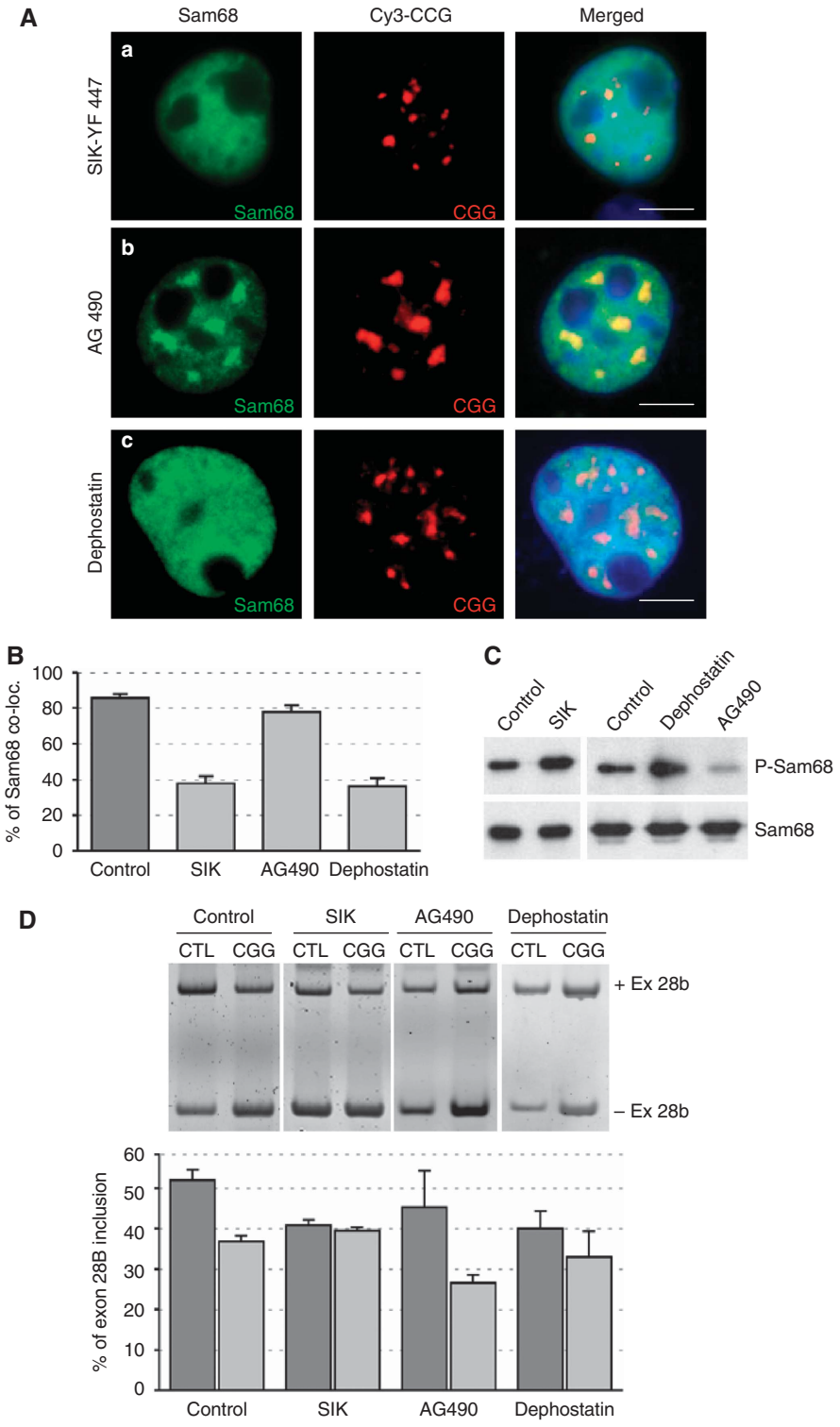


Figure 9 Tyrosine phosphorylation reduces Sam68 colocalization within CGG aggregates. (**A**) COS7 cells were transfected with a plasmid expressing 60 CGG repeats and either co-transfected with a plasmid expressing a constitutively active YF447 mutant of the SIK/Brk kinase (a), or treated with 10 μM of tyrphostin/AG490 (b) or 20 μM dephostatin (c) and analysed by FISH/IF 24 h after transfection. (**B**) The percentage of endogenous Sam68 colocalized within CGG RNA aggregates in transfected COS7 cells 24 h after transfection. (**C**) Similar to panel A but phosphorylation of Sam68 was assayed by Sam68 immunoprecipitation followed by western blotting using a Sam68-Y440 phospho-specific antibody. (**D**) Similar to panel A but COS7 cells were also co-transfected with an *ATP11B* minigene. Inclusion of exon-28B was quantified 24 h after transfection by RNA extraction followed by RT–PCR.

First, we tested Purα, but could only detect small quantities of Purα within the neuronal intranuclear inclusions in sections from human FXTAS brain as compared with abundant cytoplasmic labelling in the same neurons, which suggests that Purα is still available for its cellular function. This is in agreement with a recent report showing absence of Purα-positive inclusions in a FXTAS mouse model expressing 90 CGG repeats (Hashem *et al*, 2009). However, Purα was found

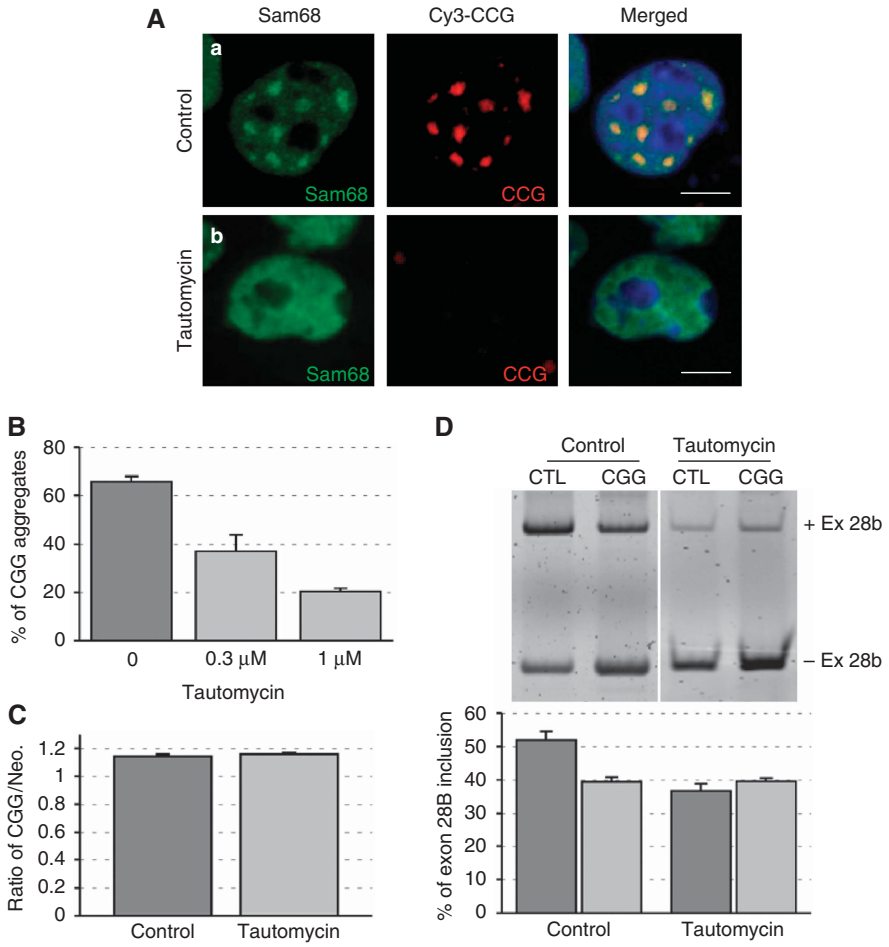


Figure 10 Tautomycin abolishes CGG and Sam68 aggregate formation. (**A**) COS7 cells were co-transfected with a plasmid expressing 60 CGG repeats and either treated with solvent (a) or 1 μM of tautomycin (b), then analysed by FISH/IF 24 h after transfection. (**B**) The percentage of cells containing CGG RNA aggregates in transfected COS7 cells treated with solvent, 0.3 or 1 μM of tautomycin. (**C**) Similar to panel A but followed by RNA extraction 24 h after transfection and qRT–PCR quantification of the CGG repeats RNA and of the neomycin mRNA. CGG and neomycin cassettes are expressed from the same pcDNA3.1 plasmid, but under different promoters. (**D**) Similar to panel A or C but COS7 cells were also co-transfected with an *ATP11B* minigene. Inclusion of exon-28B was identified 24 h after transfection by RNA extraction followed by RT–PCR.

to colocalize with HSP70 inclusions in CGG-expressing *Drosophila* and in a single reported human inclusion (Jin *et al*, 2007). A possible explanation is that Purα, which is cytoplamic, is not found for the most part in CGG aggregates, which are strictly intranuclear in the mammalian cells and FXTAS patients that we analysed (Greco *et al*, 2006). In contrast, in CGG-repeat-expressing *Drosophila*, a substantial fraction of the HSP70 inclusions were cytoplasmic and found to colocalize with Purα (Jin *et al*, 2003, 2007).

Next, we found that MBNL1, which is recruited tardily within CGG aggregates in transfected cells, also colocalizes with CGG inclusions in FXTAS patients (Supplementary Figure S5, Iwahashi *et al*, 2006). However, we found that the splicing events regulated by MBNL1 are not altered in CGG-expressing cells or in FXTAS patients (Supplementary Figure S5). These results suggest that, while MBNL1 is present within CGG aggregates, it is not immobilized and does not lose its splicing function. We also tested the presence of hnRNP-G within CGG aggregates in FXTAS patients, but none of the anti-hnRNP-G antibodies that we tested resisted the FISH conditions. Thus, whether hnRNP G colocalizes with CGG aggregates in FXTAS patients, remains to be determined.

Finally, we found that, while very late recruited within CGG aggregates, and only in dying cells, both hnRNP-A1 and hnRNP-A2/B1 were present in the CGG aggregates of FXTAS patients (data not shown; Iwahashi *et al*, 2006). By contrast, little or no hnRNP-A2/B1 aggregates were observed in CGG KI mice (Hashem *et al*, 2009 and R Willemsen, personal communication), which is consistent with smaller CGG aggregates in mice as compared with that in humans. hnRNP-A1 regulates the alternative splicing of exon-7 of the *APP* pre-mRNA (Donev *et al*, 2007), but we found no splicing alterations of *APP* in FXTAS patients (Supplementary Figure S5), suggesting that, similar to MBNL1, recruitment of hnRNP-A1 within CGG aggregates may not impair its splicing-regulatory function.

The presence of hnRNP-A1 and hnRNP-A2/B1 aggregates only in FXTAS patients is reminiscent of the ubiquitin situation, where no ubiquitin aggregates are found in CGG-transfected cells, while ubiquitin-positive inclusions are a hallmark of FXTAS patients (Greco *et al*, 2002; Arocena

et al, 2005). Overall, these data suggest that accumulation of some proteins, such as MBNL1, hnRNP-A1, hnRNP-A2 and ubiquitinylated proteins, in CGG aggregates of brain from FXTAS patients is a very late-onset event, which may not result in sequestration and loss of function of these proteins.

Sam68 nucleation model

In contrast to late recruited proteins, we found that Sam68 is early and constantly associated with CGG aggregates. Furthermore, we found that depletion of Sam68 abolished the recruitment of MBNL1 and hnRNP-G within CGG aggregates and reduced the formation of giant aggregates. These results suggest that Sam68 is essential for recruitment of novel proteins and the continuous enlargement of the CGG aggregates. Sam68 is involved in various aspects of mRNA metabolism and contains several domains enabling protein–protein interactions, allowing Sam68 to interact with various RNA-binding proteins, including hnRNP-G, hnRNP-A1, Tra2β (SFRS10), Slm1 and Slm2 (T-Star) (Venables *et al*, 1999, 2000; Paronetto *et al*, 2007). Co-transfection of these candidates with expanded CGG repeats showed that Slm1, Slm2 and Tra2β colocalized with large CGG aggregates 48 h after transfection. We propose that these proteins can in turn associate with other proteins such as MBNL1 and probably many others, ultimately resulting in massive protein aggregation disastrous for normal cell function and viability (Supplementary Figure S6B). This 'sequential recruitment' model is consistent with the late colocalization of hnRNP-A, hnRNP-G and MBNL1 in transfected cells, and with the large intranuclear inclusions, cell death and global cerebral and cerebellar atrophy observed in FXTAS patients (Greco *et al*, 2002, 2006; Tassone *et al*, 2004).

Such a model of sequential protein recruitment within CGG aggregates implies a founding and nucleating RNA–protein interaction event that would subsequently trap other proteins through indirect RNA–protein or protein–protein interactions. We propose that Sam68 is part of this founding event, as its depletion by shRNA inhibits subsequent protein aggregation and suppresses the formation of giant CGG aggregates. However, Sam68 does not bind directly to CGG RNA (data not shown), and its localization within CGG aggregates requires its N-terminal domain (Supplementary Figure S3), which contains a number of protein–protein interaction domains. These data suggest that association of Sam68 with CGG repeats it not direct but requires an intermediary RNA-binding protein (Supplementary Figure S6B). Characterization of that protein is ongoing (C Sellier, in preparation), but beyond the scope of this paper.

Sam68 protein is partially sequestered and functionally inactivated in FXTAS patients

The RNA gain-of-function model for FXTAS predicts that Sam68 protein should be sequestered by CGG RNA repeats and consequently lose its activity. FRAP analysis showed that there is a significant decrease in the mobile fraction of Sam68 in CGG-transfected cells, suggesting that a fraction of nuclear Sam68 protein is immobilized within Sam68 aggregates. Interestingly, the proportion of Sam68 immobilized in CGG-transfected cells is identical to the fraction of MBNL1 immobilized in CUG-transfected cells (Ho *et al*, 2005). Through analogy with DM, this shows partial depletion of free Sam68 in CGG-expressing cells, and hence confirms an RNA gain-of-function model for FXTAS.

Sam68 is involved in various aspects of mRNA metabolism, such as alternative splicing regulation, nuclear export, somatodendritic transport, polyadenylation and translation (reviewed by Lukong and Richard, 2003). siRNA depletion of Sam68 impairs neuronal differentiation in cell cultures (Chawla *et al*, 2009), and its ablation in mouse leads to motor coordination defects (Lukong and Richard, 2008), suggesting that partial Sam68 sequestration and loss of function might be involved in the progressive intention tremor and gait ataxia observed in FXTAS patients.

One of the important functions of Sam68 is in alternative splicing regulation (Stoss *et al*, 2001; Paronetto *et al*, 2007; Chawla *et al*, 2009). We found mis-regulation of pre-mRNA alternative splicing controlled by Sam68 in CGG-transfected cells and in FXTAS patients. Notably, analysis of alternative splicing of *ATP11B* pre-mRNA showed a splicing mis-regulation similar in FXTAS patients and in Sam68-depleted cells by shRNA transfection. These results are consistent with partial loss of function of Sam68 in CGG-transfected cells and in FXTAS patients. Phospho-type-4 ATPase-11B (*ATP11B*) is a putative flippase predicted to catalyse aminophospholipid transport and create lipid asymmetry in late secretory and endocytic compartments. Exclusion of exon-28B results in a protein isoform with a distinct C-terminal part. The physiological functions of this splice form remain to be determined.

Next, we tested other splicing events regulated by Sam68 (Chawla *et al*, 2009). However, we found no significant alterations of alternative splicing of the *KTN1*, *BIN1*, *DNCIC2*, *CLASP2* and *SGCE2* pre-mRNAs in brain samples of FXTAS patients, possibly due to differences between human patients and murine cells where they were identified as Sam68 targets (Chawla *et al*, 2009). Furthermore, we found no significant alterations of *ATP11B*, *BIN1*, *DNCIC2* and *CLASP2* alternative splicing in brain samples of Sam68-knockout mice, probably due to the compensatory effects of Sam68 paralogues, Slm1 and Slm2, which are highly expressed in brain (Stoss *et al*, 2004). We also tested the alternative splicing of *Bcl-x* and *CD44* pre-mRNAs in FXTAS patients (Matter *et al*, 2002; Paronetto *et al*, 2007). However, transcripts containing the exon-v5 of the *CD44*, or corresponding to the *Bcl-x* Short isoform, were not expressed in the brain samples that we tested.

Finally, we tested the alternative splicing of the *SMN1* and *SMN2* pre-mRNAs. Loss of the *SMN1* gene is responsible for spinal muscular atrophy (SMA), the most common inherited motor neuron disease. The near-identical *SMN2* gene does not compensate the deficiency of *SMN1* due to skipping of *SMN2* exon-7 through alternative splicing. We found that Sam68 weakly regulates the alternative splicing of an *SMN2* minigene and inhibits the inclusion of exon-7. However, we found in FXTAS patients decreased expression of *SMN2* exon-7 was, which is contradictory with depletion of Sam68 function in FXTAS. *SMN2* exon-7 inclusion is robustly stimulated by hnRNP G and Tra2β (Hofmann and Wirth, 2002; Heinrich *et al*, 2009), which are both late recruited within CGG aggregates in COS7-transfected cells. Whether hnRNP-G and Tra2β are sequestered within CGG aggregates and lose their splicing functions in FXTAS patients remain to be determined. Furthermore, whether decrease of *SMN2* exon-7 in FXTAS patients results in lower quantity of SMN protein

and is involved in the neuronal death observed in these patients remains also to be determined.

Tautomycin inhibits the formation of CGG aggregates

Sam68 activity is regulated by phosphorylation (Derry *et al*, 2000; Haegebarth *et al*, 2004; Lukong *et al*, 2005), a characteristic that made Sam68 the prototype of the Signal Transducers and Activators of RNA (STAR) family, which transduces information from signalling pathways to mRNA metabolism. Consistent with such a function, we found that tyrosine phosphorylation of Sam68 reduces its recruitment within CGG aggregates and, consequently, the deleterious effects of the CGG repeats on splicing regulation (Supplementary Figure S6A). By contrast, we found that activation of the MAP serine/threonine kinase pathway stimulates Sam68 recruitment within CGG aggregates. However, we failed to detect a direct phosphorylation of Sam68 due to the poor quality of the serine and threonine antibodies that we tested (Supplementary Figure S7).

Strikingly, among the various drugs tested, we found one, tautomycin, which not only prevents Sam68 colocalization with CGG aggregates, but also reduces CGG aggregate formation. We found that tautomycin prevents the deleterious effects of the CGG repeats on alternative splicing regulation, but due to the inherent toxicity of tautomycin, we were not able to assess whether tautomycin also reduces the toxicity of the CGG repeats. By contrast to tautomycin, depletion of Sam68 by shRNA does not alter the formation of CGG aggregates, suggesting that tautomycin may act upstream from Sam68, maybe on the uncharacterized protein that bridges CGG repeats and Sam68 (Supplementary Figure S6B). Tautomycin was first described as an inhibitor of the serine/threonine PP1 phosphatase, and more recently as an inhibitor of the glycogen synthase kinase-3β (GSK-3β) (Adler *et al*, 2009) and of the Raf1 pathway (Pinchot *et al*, 2009). Whether formation of CGG aggregates requires the PP1, GSK-3β or Raf pathway remains to be determined.

In conclusion, our results support an RNA gain-of-function mechanism in which Sam68 is partially sequestered within CGG aggregates and consequently loses its regulatory function in neurons from FXTAS patients. Furthermore, our data suggest that CGG aggregates can be dispersed, thus, bringing hope of drug treatments able to reduce CGG aggregate formation in FXTAS patients.

Materials and methods

Plasmids and constructions

Plasmids expressing 20, 40 or 60 CGG repeats were constructed by ligation of oligonucleotides containing 20 CGG repeats in pcDNA3.1. The plasmid expressing 98 CGG repeats was described previously (Arocena *et al*, 2005). *ATP11B* exon-28B was amplified using primers Fwd: 5′-AAAAAAAACAATTGCCCTAAATCTTGGTGG CAAATG and Rev: 5′-AAAAAAAACAATTGGTTGTGAGAATATCTT CACAGC, using BAC RP11-36G17 as template, and cloned within the vector pXJ41. Sam68 WT and mutants, Slm1, Slm2 and hnRNP-G expression plasmids were described previously (Venables *et al*, 2004).

Cell cultures, transfections and treatments

COS7 cells were cultured in Dulbecco's modified Eagle's Medium (DMEM), 10% foetal bovine serum and gentamicin at 37°C in 5% CO_2. PC12 were cultured in DMEM, 10% horse serum, 5% foetal calf serum and penicillin at 37°C, 5% CO_2. Cells were plated on glass coverslips in a 24-well plate for immunofluorescence and transfected 24 h after plating in DMEM + 0.1% foetal bovine serum to block cell divisions, using either FugenHD (Roche) for COS7 cells or Lipofectamine 2000 (Invitrogen) for PC12 cells. PC12 cells were differentiated 6 h after transfection by growing cells in 1% horse serum, 1% foetal calf serum plus 50 ng/ml of NGF (Clinisciences). Primary cultures of hippocampal neurons were obtained from WT E18 mouse and grown into 24-wells plates in 500 μl neurobasal medium (Gibco), 1 × B27 (Gibco), 0.5 mM L-glutamine and penicillin at 37°C, 5% CO_2, and were transfected after 4 DIV with Lipofectamine 2000 (Invitrogen).

Cell treatments

Cells were incubated 6 h after transfection in 50 μM PD98059, 4 nM okadaic acid, 40 ng/ml TPA, 10 μM AG490, 20 μM dephostatin or 1 μM tautomycin (Calbiochem).

Patients and brain sections

FXTAS patients have been described previously (case 6, 7 9 and 10 of Greco *et al*, 2006). KI mouse (72 weeks old, 98 CGG repeats) or human brain sections were deparaffinized two times for 20 min in Histosol Plus (Shandon) and dehydrated as follows: twice in ethanol 100% (5 min), twice in ethanol 95% (5 min), once in ethanol 80% (5 min), once in ethanol 70% (5 min) and rinsed in TBS-Tween 1% before FISH.

FISH combined with immunofluorescence

Glass coverslips containing plated cells or brain sections on slides treated as described above were fixed in 4% paraformaldehyde in PBS (pH 7.4) for 15 min and washed three times with PBS. The coverslips were incubated for 5 min in PBS/0.5% Triton X-100 and washed three times with PBS before pre-hybridization in 40% DMSO, 40% formamide, 10% BSA (10 mg/ml), 2 × SCC for 30 min. The coverslips were hybridized for 2 h in 40% formamide, 10% DMSO, 2 × SCC, 2 mM vanadyl ribonucleoside, 60 μg/ml tRNA, 30 μg/ml BSA plus 0.75 μg $(CCG)_{8x}$-Cy3 DNA oligonucleotide probe (Sigma). The coverslips were washed twice in 2 × SCC/50% formamide and twice in 2 × SCC. Following FISH, the coverslips were washed twice successively in 2 × SCC/50% formamide, in 2 × SCC and in PBS. The coverslips were incubated overnight with primary anti-hnRNP-G (1/200 dilution; Heinrich *et al*, 2009), hnRNP-A2/B1 (1/200 dilution, clone DP3B3; Tebu), hnRNP-A1 (1/200 dilution, clone 9H10; Abcam), MBNL1 (1/200 dilution, clone HL1822 3A4-1E9; Sigma) and Sam68 (1/400 dilution, C20 SC-333; Santa Cruz Biotechnology) at 4°C. The coverslips were washed twice with PBS before incubation with a goat anti-rabbit secondary antibody conjugated with Alexa-Fluor 488 (1/500 dilution; Fisher) or Cy5 (1/500 dilution; Interchim) for 60 min. Then, the coverslips were incubated for 10 min in 2 × SCC/DAPI (1/10 000 dilution) and rinsed twice in 2 × SSC before mounting in Pro-Long media (Molecular Probes). Slides were examined using either a simple fluorescence microscope (Leica) or a Leica DM4000 B confocal microscope, equipped with a Leica 100 × HCX Plan Apo CS 1.40 objective, in 1-μm optical sections.

FRAP analysis

COS7 cells were plated in 35-mm glass base dish (Iwaki) and transfected with Fugene HD 24 h after plating. Twenty-four hours after transfection, FRAP experiments were performed using a Leica DM4000 B confocal microscope combined to a heated stage. The cells were maintained for 15 min in growth media at 37°C with no CO_2 or humidifier systems; five single scans were obtained followed by five bleach pulses. After photobleaching, images were taken every 10 s for 150 s (post-bleach 1) followed by 40 acquisitions every 20 s (post-bleach 2). Fluorescence intensities were calculated using the Image J software and normalized by total cellular fluorescence intensity. As foci are mobile, an area around the foci was determined for the duration of the experiment.

MALDI analysis

Nuclear extracts were prepared from COS7 cells or mouse brain extract as described by Dignam *et al* (1983). Nuclear extract was passed over a CGG_{60x} *in vitro* T7 transcribed and biotinylated RNA (Ambion) bound to a streptavidin agarose column (Invitrogen) in the presence of KCl (100 mM), HEPES (10 mM) and $MgCl_2$ (1 mM). The column was washed with 33 mM $MgCl_2$ and glacial acetic acid. The eluted protein was separated by gel electrophoresis and detected by silver staining. The protein bands were excised,

digested and identified using a Reflex IV MALDI-TOF spectrometer (Bruker Daltonics) and the Profound search engine, as described by Argentini *et al*, 2008.

RNA isolation and PCR

COS cells were cultured as described above in a six-well plate. RNA was isolated 24 h after transfection using a GenElute kit (Sigma). Endogenous *ATP11B* splicing analysis was performed on total RNA extracted (Trizol reagent) from control or FXTAS patient brain samples. cDNA synthesis reactions were performed with Superscript II (Invitrogen). The primers are described in Supplementary Table 4. PCRs were performed with *Taq* polymerase (Roche). The conditions for *SMN2* and *Bcl-x* minigenes were described previously (Paronetto *et al*, 2007; Heinrich *et al*, 2009). The conditions for *ATP11B* minigene were 4 min at 94°C; 30 cycles of 40 s at 94°C, 45 s at 60°C and 1 min at 72°C; and a final extension at 72°C for 4 min. The PCR reaction products were analysed on 8% polyacrylamide gels.

Statistical analysis

The percentage of endogenous Sam68, MBNL1 and hnRNP-G colocalized within CGG RNA is expressed as the number of nuclei presenting colocalization/total number of nuclei containing CGG aggregates. Three independent transfections totalling a hundred cells were counted. Results are presented as mean ± s.d. The balance between the mRNA levels of the *SMN2*, *Bcl-x* and *ATP11B* splicing variants is calculated as $((\text{mRNA}+\text{exon})/(\text{mRNA}-\text{exon}+\text{mRNA}+\text{exon}))\times 100$. The results are derived from at least three independent experiments. The error bars in the figures indicate the s.e.m.'s. Statistical differences were calculated using *t*-test.

Supplementary data

Supplementary data are available at *The EMBO Journal* Online (http://www.embojournal.org).

Acknowledgements

We thank the imagery, cell culture and MALDI platforms of the IGBMC; Thomas Cooper for the gift of the DT960 and *INSR* minigene plasmid; Laura Ranum for the gift of the CCTG300 expression plasmid; Charles Thornton for the gift of the polyclonal MBNL1 antibody; Angela Tyner for the gift of the *SIK/Brk* constructs; Claudio Sette for the gift of the *Bcl-x*, shSam68 and *Sam68* WT and mutant constructs; Stefan Stamm for the gift of the *SMN2* minigene; Stefan Hubner for the gift of the GFP-lamin-A construct; and all members of the French DM Network for fruitful discussion. This work was supported by INSERM AVENIR funding (NCB), ANR GENOPAT grant P007942 (NCB), NIH-NINDS grant NS062411 (RW), NIH Roadmap Initiative DE019583 and AG032119 (PJH), BBSRC grant BB/D013917/1 (DJE), Wellcome Trust grant number WT080368MA (DJE) and an ORSAS international studentship (YL).

Conflict of interest

The authors declare that they have no conflict of interest.

References

Adler JT, Cook M, Luo Y, Pitt SC, Ju J, Li W, Shen B, Kunnimalaiyaan M, Chen H (2009) Tautomycetin and tautomycin suppress the growth of medullary thyroid cancer cells via inhibition of glycogen synthase kinase-3beta. *Mol Cancer Ther* **8:** 914–920

Argentini M, Strub JM, Carapito C, Sanglier S, Van-Dorsselaer A (2008) An optimized MALDI mass spectrometry method for improved detection of lysine/arginine/histidine free peptides. *J Proteome Res* **7:** 5062–5069

Arocena DG, Iwahashi CK, Won N, Beilina A, Ludwig AL, Tassone F, Schwartz PH, Hagerman PJ (2005) Induction of inclusion formation and disruption of lamin A/C structure by premutation CGG repeat RNA in human cultured neural cells. *Human Mol Genet* **14:** 3661–3671

Brouwer JR, Huizer K, Severijnen LA, Hukema RK, Berman RF, Oostra BA, Willemsen R (2008) CGG repeat length and neuropathological and molecular correlates in a mouse model for fragile X-associated tremor/ataxia syndrome. *J Neurochem* **107:** 1671–1682

Chawla G, Lin CH, Han A, Shiue L, Ares Jr M, Black DL (2009) Sam68 regulates a set of alternatively spliced exons during neurogenesis. *Mol Cell Biol* **29:** 201–213

Clark TA, Schweitzer AC, Chen TX, Staples MK, Lu G, Wang H, Williams A, Blume JE (2007) Discovery of tissue-specific exons using comprehensive human exon microarrays. *Genome Biol* **8:** 64

Derry JJ, Richard S, Valderrama Carvajal H, Ye X, Vasioukhin V, Cochrane AW, Chen T, Tyner AL (2000) SIK (BRK) phosphorylates Sam68 in the nucleus and negatively regulates its RNA binding ability. *Mol Cell Biol* **20:** 6114–6126

Dignam JD, Lebovitz RM, Roeder RG (1983) Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res* **11:** 1475–1489

Donev R, Newall A, Thome J, Sheer D (2007) A role for SC35 and hnRNPA1 in the determination of amyloid precursor protein isoforms. *Mol Psychiatry* **12:** 681–690

Greco CM, Berman RF, Martin RM, Tassone F, Schwartz PH, Chang A, Trapp BD, Iwahashi C, Brunberg J, Grigsby J, Hessl D, Becker EJ, Papazian J, Leehey MA, Hagerman RJ, Hagerman PJ (2006) Neuropathology of fragile X-associated tremor/ataxia syndrome (FXTAS). *Brain* **1:** 243–255

Greco CM, Hagerman RJ, Tassone F, Chudley AE, Del Bigio MR, Jacquemont S, Leehey M, Hagerman PJ (2002) Neuronal intranuclear inclusions in a new cerebellar tremor/ataxia syndrome among fragile X carriers. *Brain* **125:** 1760–1771

Haegebarth A, Heap D, Bie W, Derry JJ, Richard S, Tyner AL (2004) The nuclear tyrosine kinase BRK/SIK phosphorylates and inhibits the RNA-binding activities of the Sam68-like mammalian proteins SLM-1 and SLM-2. *J Biol Chem* **279:** 54398–54404

Hagerman PJ, Hagerman RJ (2004) The fragile-X premutation: a maturing perspective. *Am J Hum Genet* **74:** 805–816

Hagerman RJ, Leehey M, Heinrichs W, Tassone F, Wilson R, Hills J, Grigsby J, Gage B, Hagerman PJ (2001) Intention tremor, parkinsonism, and generalized brain atrophy in male carriers of fragile X. *Neurology* **1:** 127–130

Hashem V, Galloway JN, Mori M, Willemsen R, Oostra BA, Paylor R, Nelson DL (2009) Ectopic expression of CGG containing mRNA is neurotoxic in mammals. *Hum Mol Genet* **18:** 2443–2451

Heinrich B, Zhang Z, Raitskin O, Hiller M, Benderska N, Hartmann AM, Bracco L, Elliott D, Ben-Ari S, Soreq H, Sperling J, Sperling R, Stamm S (2009) Heterogeneous nuclear ribonucleoprotein G regulates splice site selection by binding to CC(A/C)-rich regions in pre-mRNA. *J Biol Chem* **284:** 14303–14315

Ho TH, Savkur RS, Poulos MG, Mancini MA, Swanson MS, Cooper TA (2005) Colocalization of muscleblind with RNA foci is separable from mis-regulation of alternative splicing in myotonic dystrophy. *J Cell Sci* **118** (Part 13)**:** 2923–2933

Hofmann Y, Wirth B (2002) hnRNP-G promotes exon 7 inclusion of survival motor neuron (SMN) via direct interaction with Htra2-beta1. *Hum Mol Genet* **11:** 2037–2049

Iwahashi CK, Yasui DH, An HJ, Greco CM, Tassone F, Nannen K, Babineau B, Lebrilla CB, Hagerman RJ, Hagerman PJ (2006) Protein composition of the intranuclear inclusions of FXTAS. *Brain* **129:** 256–271

Jacquemont S, Hagerman RJ, Leehey M, Grigsby J, Zhang L, Brunberg JA, Greco C, Des Portes V, Jardini T, Levine R, Berry-Kravis E, Brown WT, Schaeffer S, Kissel J, Tassone F, Hagerman PJ (2003) Fragile X premutation tremor/ataxia syndrome: molecular, clinical, and neuroimaging correlates. *Am J Hum Genet* **4:** 869–878

Jacquemont S, Hagerman RJ, Leehey MA, Hall DA, Levine RA, Brunberg JA, Zhang L, Jardini T, Gane LW, Harris SW, Herman K, Grigsby J, Greco CM, Berry-Kravis E, Tassone F, Hagerman PJ (2004) Penetrance of the fragile X-associated tremor/ataxia syndrome in a premutation carrier population. *JAMA* **4:** 460–469

Jin P, Duan R, Qurashi A, Qin Y, Tian D, Rosser TC, Liu H, Feng Y, Warren ST (2007) Pur alpha binds to rCGG repeats and modulates repeat-mediated neurodegeneration in a *Drosophila* model of fragile X tremor/ataxia syndrome. *Neuron* **55:** 556–564

Jin P, Zarnescu DC, Zhang F, Pearson CE, Lucchesi JC, Moses K, Warren ST (2003) RNA-mediated neurodegeneration caused by the fragile X premutation rCGG repeats in *Drosophila*. *Neuron* **39:** 739–747

Kenneson A, Zhang F, Hagedorn CH, Warren ST (2001) Reduced FMRP and increased FMR1 transcription is proportionally associated with CGG repeat number in intermediate-length and premutation carriers. *Hum Mol Genet* **10:** 1449–1454

Leehey MA, Berry-Kravis E, Goetz CG, Zhang L, Hall DA, Li L, Rice CD, Lara R, Cogswell J, Reynolds A, Gane L, Jacquemont S, Tassone F, Grigsby J, Hagerman RJ, Hagerman PJ (2008) FMR1 CGG repeat length predicts motor dysfunction in premutation carriers. *Neurology* **16:** 1397–1402

Liu Y, Bourgeois C, Pang S, Kudla M, Dreumont N, Kister L, Sun YH, Stevenin J, Elliott DJ (2009) The germ cell nuclear proteins hnRNP G-T and RBMY activate a testis-specific exon. *PLOS Genet* **5:** e1000707

Lukong KE, Larocque D, Tyner AL, Richard S (2005) Tyrosine phosphorylation of Sam68 by breast tumor kinase regulates intranuclear localization and cell cycle progression. *J Biol Chem* **280:** 38639–38647

Lukong KE, Richard S (2003) Sam68, the KH domain-containing superstar. *Biochim Biophys Acta* **1653:** 73–86

Lukong KE, Richard S (2008) Motor coordination defects in mice deficient for the Sam68 RNA-binding protein. *Behav Brain Res* **189:** 357–363

Matter N, Herrlich P, König H (2002) Signal-dependent regulation of splicing via phosphorylation of Sam68. *Nature* **420:** 691–695

Mermoud JE, Cohen P, Lamond AI (1992) Ser/Thr-specific protein phosphatases are required for both catalytic steps of pre-mRNA splicing. *Nucleic Acids Res* **20:** 5263–5269

Novoyatleva T, Heinrich B, Tang Y, Benderska N, Butchbach ME, Lorson CL, Lorson MA, Ben-Dov C, Fehlbaum P, Bracco L, Burghes AH, Bollen M, Stamm S (2008) Protein phosphatase 1 binds to the RNA recognition motif of several splicing factors and regulates alternative pre-mRNA processing. *Hum Mol Genet* **17:** 52–70

Ostra BA, Willemesen R (2009) FMR1: a gene with three faces. *Biochim Biophys Acta* **790:** 467–477

Paronetto MP, Achsel T, Massiello A, Chalfant CE, Sette C (2007) The RNA-binding protein Sam68 modulates the alternative splicing of Bcl-x. *J Cell Biol* **176:** 929–939

Pinchot SN, Adler JT, Luo Y, Ju J, Li W, Shen B, Kunnimalaiyaan M, Chen H (2009) Tautomycin suppresses growth and neuroendocrine hormone markers in carcinoid cells through activation of the Raf-1 pathway. *Am J Surg* **197:** 313–319

Primerano B, Tassone F, Hagerman RJ, Hagerman P, Amaldi F, Bagni C (2002) Reduced FMR1 mRNA translation efficiency in fragile X patients with premutations. *RNA* **12:** 1482–1488

Ranum LP, Cooper TA (2006) RNA-mediated neuromuscular disorders. *Annu Rev Neurosci* **29:** 259–277

Sofola OA, Jin P, Qin Y, Duan R, Liu H, de Haro M, Nelson DL, Botas J (2007) RNA-binding proteins hnRNP A2/B1 and CUGBP1 suppress fragile X CGG premutation repeat-induced neurodegeneration in a *Drosophila* model of FXTAS. *Neuron* **55:** 565–571

Stoss O, Novoyatleva T, Gencheva M, Olbrich M, Benderska N, Stamm S (2004) p59(fyn)-mediated phosphorylation regulates the activity of the tissue-specific splicing factor rSLM-1. *Mol Cell Neurosci* **27:** 8–21

Stoss O, Olbrich M, Hartmann AM, Konig H, Memmott J, Andreadis A, Stamm S (2001) The STAR/GSG family protein rSLM-2 regulates the selection of alternative splice sites. *J Biol Chem* **276:** 8665–8673

Tassone F, Adams J, Berry-Kravis EM, Cohen SS, Brusco A, Leehey MA, Li L, Hagerman RJ, Hagerman PJ (2007) CGG repeat length correlates with age of onset of motor signs of the fragile X-associated tremor/ataxia syndrome (FXTAS). *Am J Med Genet B Neuropsychiatr Genet* **4:** 566–569

Tassone F, Hagerman RJ, Chamberlain WD, Hagerman PJ (2000a) Transcription of the FMR1 gene in individuals with fragile X syndrome. *Am J Med Genet* **97:** 195–203

Tassone F, Hagerman RJ, Loesch DZ, Lachiewicz A, Taylor AK, Hagerman PJ (2000b) Fragile X males with unmethylated, full mutation trinucleotide repeat expansions have elevated levels of FMR1 messenger RNA. *Am J Med Genet* **94:** 232–236

Tassone F, Iwahashi C, Hagerman PJ (2004) FMR1 RNA within the intranuclear inclusions of fragile X-associated tremor/ataxia syndrome (FXTAS). *RNA Biol* **1:** 103–105

Van Dam D, Errijgers V, Kooy RF, Willemsen R, Mientjes E, Oostra BA, De Deyn PP (2005) Cognitive decline, neuromotor and behavioural disturbances in a mouse model for fragile-X-associated tremor/ataxia syndrome (FXTAS). *Behav Brain Res* **162:** 233–239

Venables JP, Dalgliesh C, Paronetto MP, Skitt L, Thornton JK, Saunders PT, Sette C, Jones KT, Elliott DJ (2004) SIAH1 targets the alternative splicing factor T-STAR for degradation by the proteasome. *Hum Mol Genet* **13:** 1525–1534

Venables JP, Elliott DJ, Makarova OV, Makarov EM, Cooke HJ, Eperon IC (2000) RBMY, a probable human spermatogenesis factor, and other hnRNP G proteins interact with Tra2beta and affect splicing. *Hum Mol Genet* **9:** 685–694

Venables JP, Vernet C, Chew SL, Elliott DJ, Cowmeadow RB, Wu J, Cooke HJ, Artzt K, Eperon IC (1999) TSTAR/ÉTOILE: a novel relative of SAM68 that interacts with an RNA-binding protein implicated in spermatogenesis. *Hum Mol Genet* **8:** 959–969

Wheeler TM, Thornton CA (2007) Myotonic dystrophy: RNA-mediated muscle disease. *Curr Opin Neurol* **20:** 572–576

Willemsen R, Hoogeveen-Westerveld M, Reis S, Holstege J, Severijnen LA, Nieuwenhuizen IM, Schrier M, van Unen L, Tassone F, Hoogeveen AT, Hagerman PJ, Mientjes EJ, Oostra BA (2003) The FMR1 CGG repeat mouse displays ubiquitin-positive intranuclear neuronal inclusions; implications for the cerebellar tremor/ataxia syndrome. *Hum Mol Genet* **12:** 949–959

Cell Reports
Article

Sequestration of DROSHA and DGCR8 by Expanded CGG RNA Repeats Alters MicroRNA Processing in Fragile X-Associated Tremor/Ataxia Syndrome

Chantal Sellier,[1,2] Fernande Freyermuth,[1] Ricardos Tabet,[1] Tuan Tran,[3] Fang He,[4] Frank Ruffenach,[1] Violaine Alunni,[1] Herve Moine,[1] Christelle Thibault,[1] Adeline Page,[1] Flora Tassone,[5] Rob Willemsen,[7] Matthew D. Disney,[3] Paul J. Hagerman,[5,6] Peter K. Todd,[4] and Nicolas Charlet-Berguerand[1,*]

[1]Department of Translational Medicine, IGBMC, Illkirch 67400, France
[2]College de France, Paris 75000, France
[3]Department of Chemistry, The Scripps Research Institute, Jupiter, FL 33458, USA
[4]Department of Neurology, University of Michigan, Ann Arbor, MI 48109, USA
[5]M.I.N.D. Institute
[6]Department of Biochemistry and Molecular Medicine
University of California, Davis, Sacramento, CA 95817, USA
[7]Department of Clinical Genetics, Erasmus MC, Rotterdam 3000 DR, The Netherlands
*Correspondence: ncharlet@igbmc.fr
http://dx.doi.org/10.1016/j.celrep.2013.02.004

SUMMARY

Fragile X-associated tremor/ataxia syndrome (FXTAS) is an inherited neurodegenerative disorder caused by the expansion of 55–200 CGG repeats in the 5′ UTR of *FMR1*. These expanded CGG repeats are transcribed and accumulate in nuclear RNA aggregates that sequester one or more RNA-binding proteins, thus impairing their functions. Here, we have identified that the double-stranded RNA-binding protein DGCR8 binds to expanded CGG repeats, resulting in the partial sequestration of DGCR8 and its partner, DROSHA, within CGG RNA aggregates. Consequently, the processing of microRNAs (miRNAs) is reduced, resulting in decreased levels of mature miRNAs in neuronal cells expressing expanded CGG repeats and in brain tissue from patients with FXTAS. Finally, overexpression of DGCR8 rescues the neuronal cell death induced by expression of expanded CGG repeats. These results support a model in which a human neurodegenerative disease originates from the alteration, in *trans*, of the miRNA-processing machinery.

INTRODUCTION

Fragile X-associated tremor/ataxia syndrome (FXTAS) is a neurodegenerative disorder that affects older adult males who are carriers of an expansion of 55–200 CGG repeats in the 5′ UTR of the fragile X mental retardation 1 (*FMR1*) gene (Hagerman et al., 2001). The clinical features of FXTAS include progressive intention tremor and gait ataxia, frequently accompanied by progressive cognitive decline, parkinsonism, peripheral neuropathy, and autonomic dysfunctions (Jacquemont et al., 2003). Principal neuropathology of FXTAS includes mild brain atrophy and white matter lesions with the presence of ubiquitin-positive nuclear neuronal and astrocytic inclusions (Greco et al., 2002), which contain the expanded CGG RNA repeats (Tassone et al., 2004). In contrast to fragile X syndrome, where full mutations (>200 CGG repeats) result in hypermethylation and silencing of the *FMR1* gene, FXTAS carriers of shorter CGG expansions (55–200 CGG repeats) present increased expression of *FMR1* mRNA levels and normal, or near-normal, FMRP expression (Tassone et al., 2000). These observations suggest a toxic RNA gain-of-function model for FXTAS. In support of that model, cellular and transgenic *Drosophila* and mouse models demonstrate that the sole expression of a mutant RNA containing expanded CGG repeats is sufficient to induce the formation of ubiquitin-positive aggregates and to cause a pathology similar to human FXTAS (Willemsen et al., 2003; Jin et al., 2003; Arocena et al., 2005; Entezam et al., 2007; Hashem et al., 2009). A toxic RNA gain-of-function model predicts that expanded CGG repeats are pathogenic by sequestering specific RNA-binding proteins, resulting in loss of their normal functions and, ultimately, in neuronal cell dysfunction and death. Consistent with a titration model, various proteins were found to colocalize with CGG or ubiquitin-positive inclusions (Iwahashi et al., 2006; Jin et al., 2007; Sofola et al., 2007; Sellier et al., 2010); however, the pathological consequences of their recruitment are unclear, suggesting that the protein(s) sequestered within CGG RNA aggregates and responsible for the neuronal cell death, remains to be identified.

MicroRNAs (miRNAs) are small, conserved, noncoding RNAs that are key components of posttranscriptional gene regulation and are involved in the control of many fundamental processes, including both differentiation and survival of neurons (Schaefer et al., 2007; Davis et al., 2008; De Pietri Tonelli et al., 2008; Stark et al., 2008; Haramati et al., 2010; Hébert et al., 2010; Huang et al., 2010; Fénelon et al., 2011; Schofield et al., 2011). miRNAs are initially transcribed by the RNA polymerase II as primary miRNA (pri-miRNAs) transcripts, which are processed into precursor miRNAs (pre-miRNAs) by the type III RNase,

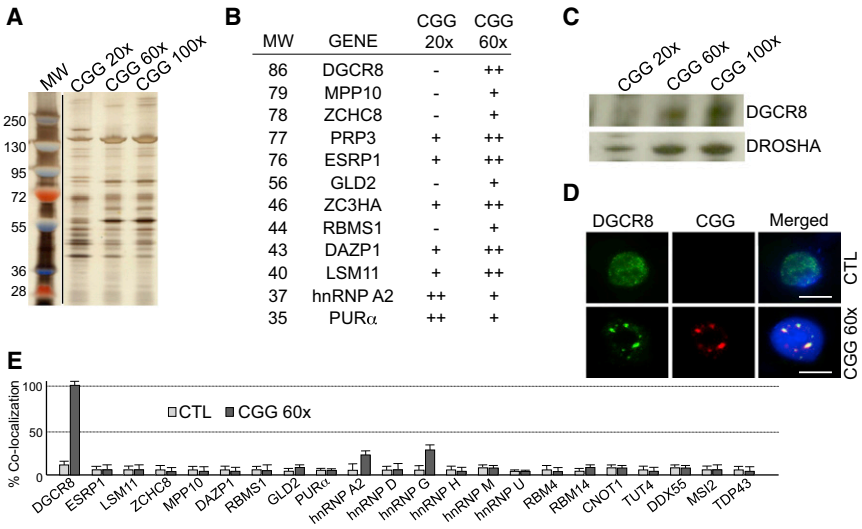


Figure 1. Identification of Proteins Associated with Expanded CGG Repeats

(A) Silver staining of proteins extracted from mouse brain and captured on streptavidin resin coupled to biotinylated in-vitro-transcribed RNA containing expanded CGG repeats of normal or pathogenic size. MW, molecular weight.

(B) List of the main proteins identified by nanoLS-MS/MS that are preferentially associated with 20 or 60 CGG repeats.

(C) Western blotting against DGCR8 or DROSHA of mouse brain proteins captured on 20 or 60 CGG RNA repeat columns.

(D) RNA FISH against CGG repeats, coupled to IF against DGCR8, of COS7 cells transfected with a plasmid expressing either no repeats (CTL) or 60 CGG repeats. Magnification, 630×. Scale bars, 10 μm.

(E) Percentage of colocalization of the endogenous-tested candidate proteins within CGG RNA aggregates analyzed by RNA FISH coupled to IF on COS7 transfected for 24 hr with a plasmid expressing either no repeats or 60 CGG repeats.

See also Tables S1, S2, and Figure S1.

DROSHA, and the double-stranded RNA-binding protein, DGCR8, which anchors DROSHA to the pri-miRNA transcript (Lee et al., 2003; Denli et al., 2004; Landthaler et al., 2004; Gregory et al., 2004; Han et al., 2004; Wang et al., 2007). Pre-miRNAs are then exported into the cytoplasm, where they are processed into mature miRNAs by the DICER enzyme. Global reduction of miRNA expression, for example through inactivation of DICER or DGCR8 in mice, results in embryonic lethality and if conditionally lost in brain, leads to neuronal dysfunction and cell death (Schaefer et al., 2007; Davis et al., 2008; De Pietri Tonelli et al., 2008; Stark et al., 2008; Haramati et al., 2010; Hébert et al., 2010; Huang et al., 2010; Fénelon et al., 2011; Schofield et al., 2011).

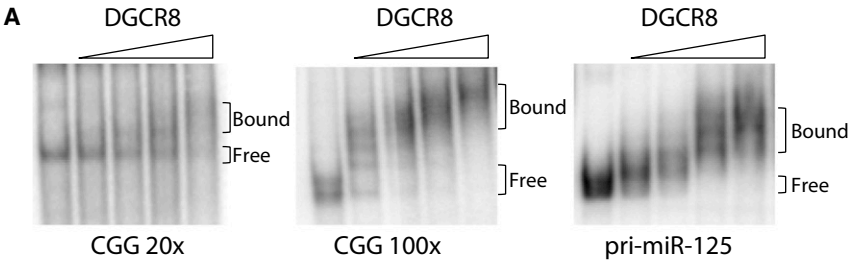
Here, we find that DGCR8 binds preferentially to expansions of CGG repeats of pathogenic length. This association results in titration of DGCR8 pri-miRNA-binding activity and in the partial sequestration of DGCR8 and its partner, DROSHA, within CGG RNA aggregates. Consequently, the processing of pri-miRNAs is reduced in cells expressing expanded CGG repeats, and in brain samples from patients with FXTAS, resulting in decreased levels of mature miRNAs. Finally, the expression of pathogenic-expanded CGG repeats in cultured mouse cortical neurons results in decreased dendritic complexity and reduced neuronal cell viability. Importantly, the sole overexpression of DGCR8 restored to normal both the dendritic morphological abnormalities and the loss of neuronal cells, demonstrating that titration of DGCR8 by expanded CGG repeats is a leading event to CGG-induced neuronal cell death.

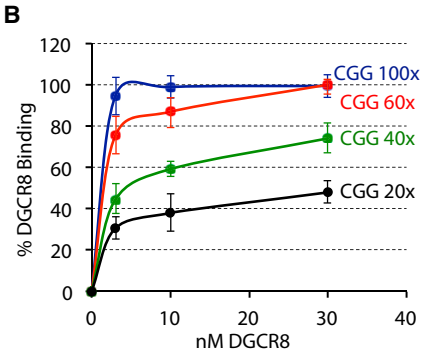
RESULTS

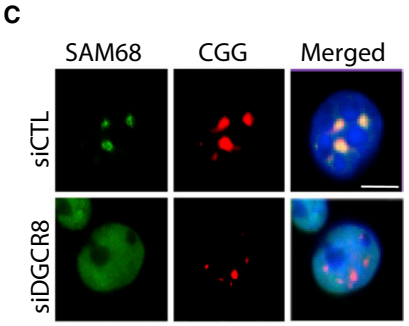
Identification of Proteins Associated with Expanded CGG Repeats

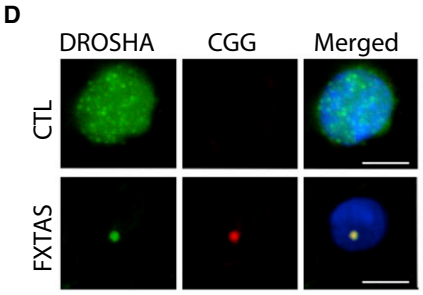
To identify proteins involved in FXTAS physiopathology, proteins extracted from mouse brain nuclei were captured on streptavidin resin coupled to biotinylated RNA composed of nonpathogenic (20 CGG repeats) or pathogenic expansions (60 or 100 CGG repeats), eluted, separated on SDS-PAGE, and identified by nano-LC-MS/MS analysis (Figure 1A). More than 30 RNA-binding proteins were identified (Figure 1B; Tables S1 and S2), including hnRNP A2/B1, hnRNP G, and PURα, which were also identified in previous studies (Iwahashi et al., 2006; Jin et al., 2007; Sofola et al., 2007; Sellier et al., 2010). However, these proteins were recruited preferentially onto short CGG repeats compared to long pathogenic CGG stretches (Figure 1B). In contrast, ten proteins (ESRP1, PRP3, ZC3HA, LSM11, ZCHC8, MPP10, DAZP1, RBMS1, GLD2, and DGCR8) were found preferentially associated with expanded CGG repeats of pathogenic size (Figure 1B). We confirmed the preferential binding of DGCR8 on long CGG stretches by western blotting on the protein eluted from the RNA affinity columns (Figure 1C). The partner of DGCR8, DROSHA, was also preferentially recruited to long expanded CGG repeats (Figure 1C).

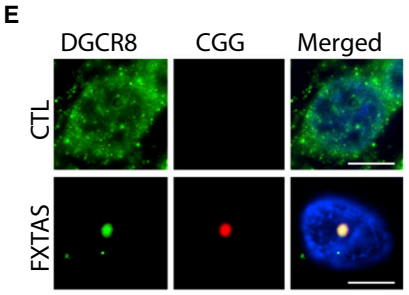
Next, to discard nonspecific RNA-binding proteins, we tested whether these candidate proteins colocalize with the aggregates formed by expanded CGG repeats in transfected COS7 cells (Figures S1A and S1B). Among the ten candidates tested, only three (PRP3, ZC3HA, and DGCR8) colocalized with CGG RNA aggregates (Figures 1D and S1A). However, we noted that PRP3 and ZC3HA were naturally localized in speckles, which are normal nuclear structures enriched in some splicing factors and that were found previously to be associated with CGG RNA aggregates in transfected cells (Sellier et al., 2010). Thus, neither PRP3 nor ZC3HA was specifically recruited within CGG RNA aggregates (Figure S1A). In contrast, DGCR8 presented a diffuse pattern within the nucleoplasm, but upon expression of 60 expanded pathogenic CGG repeats, DGCR8 changed localization and was recruited within CGG RNA inclusions (Figure 1D). As further controls, we also tested the colocalization of proteins, such as hnRNP A2/B1 and PURα, which were found mostly associated with 20 biotinylated CGG RNA repeats in vitro. Consistent with a

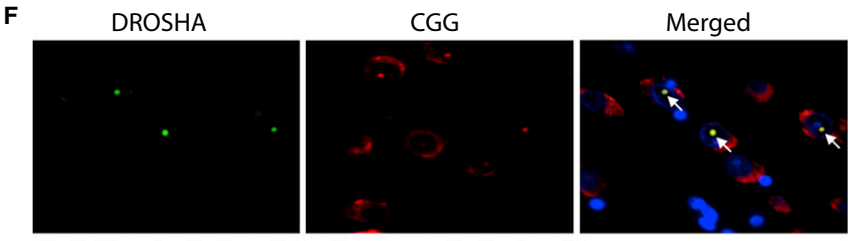












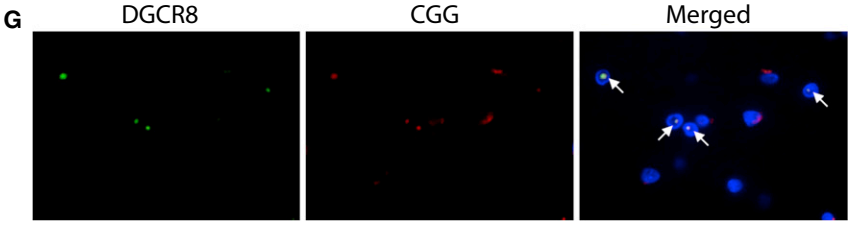


Figure 2. DGCR8 and DROSHA Bind to Expanded CGG RNA Repeats

(A) Gel shift assays of purified bacterial recombinant His-DGCR8Δ with 100 nM (30,000 cpm) of uniformly α-[^{32}P]CTP internally labeled in-vitro-transcribed RNA containing 20 or 100 CGG repeats or the pri-miR-125.

(B) Quantification of DGCR8 binding to 20, 40, 60, or 100 CGG RNA repeats.

(C) RNA FISH against CGG repeats, coupled to IF against SAM68, on COS7 cells cotransfected with a plasmid expressing 60 CGG repeats and a siRNA against the luciferase (siCTL) or against DGCR8.

(D and E) RNA FISH of CGG repeats coupled to IF of DROSHA or DGCR8 on brain sections (hippocampal area) of age-matched control or patients with FXTAS. Magnification, 630×. Nuclei were counterstained with DAPI.

(F and G) Larger fields of CGG RNA FISH coupled to IF of DROSHA or DGCR8 on brain sections of patients with FXTAS. White arrows indicate CGG RNA aggregates. Note some nonspecific perinuclear background inherent to RNA FISH of autopsied human brain samples.

Scale bars, 10 μm.

See also Figure S2.

preferential binding to normal but not to expanded CGG repeats, we found no significant colocalization of these candidates with CGG RNA aggregates at early time points (Figure 1E). Thus, we pursued our study on the one protein found specifically associated with expanded CGG repeats of pathogenic size, DGCR8.

DGCR8 Binds to Mutated RNA Containing Expanded CGG Repeats

Recruitment of DGCR8 to biotinylated CGG RNA questions whether DGCR8 binds directly, or through indirect protein-protein interactions, to expanded CGG repeats. Gel shift assays showed that purified recombinant HIS-tagged DGCR8 protein binds directly to RNAs containing expanded CGG repeats of pathogenic size (60 and 100 CGG repeats; Figures 2A, S2A, and S2B). In contrast, DGCR8 binds weakly to RNAs containing CGG stretch of nonpathogenic size, such as 20 or 40 CGG repeats (Figures 2A, S2A, and S2B). Furthermore, DGCR8 binds to expanded CGG repeats with an affinity similar to control pri-miRNAs, such as pri-miR-124, pri-miR-125, or pri-Let-7 (Figures 2A and S2A). UV-cross-linking assays confirmed that purified recombinant DGCR8 binds directly to control pri-miR-124 and pri-miR-125, as well as to expanded CGG repeats of pathogenic size, but does not bind to short CGG stretch (Figure S2B).

The in vitro binding of DGCR8 to long expanded CGG repeats prompted us to investigate whether DROSHA and DGCR8 are

recruited within the nuclear aggregates found in cell models of FXTAS. CGG RNA aggregates are dynamic nuclear structures that accumulate various proteins in a time-dependent manner (Sellier et al., 2010), which raises the question of the timing of DROSHA and DGCR8 recruitment. Analysis of the formation of CGG RNA aggregates at various time points after transfection of COS7 cells with a plasmid expressing 60 CGG repeats revealed that DROSHA and DGCR8 colocalized within CGG aggregates from the time of their formation (6–8 hr posttransfection; data not shown). Furthermore, depletion of either DROSHA or DGCR8 by siRNA reduced the recruitment of SAM68, one of the earliest proteins found to colocalize with CGG aggregates (Figures 2C and S2C). Next, coimmunoprecipitation experiments demonstrated that DROSHA and DGCR8 interact with SAM68, suggesting that the recruitment of SAM68 within CGG RNA aggregates is mediated through protein-protein interactions with DROSHA or DGCR8 (Figure S2D). These results suggest that in transfected cells, DROSHA and DGCR8 are among the first proteins to be recruited within the CGG RNA aggregates and are essential for the further aggregation of other proteins, such as SAM68. These results were obtained in cells expressing large amounts of expanded CGG RNA repeat. Thus, to rule out any overexpression bias, we tested the localization of endogenous DROSHA and DGCR8 in brain sections from patients with FXTAS. RNA FISH coupled to immunofluorescence (IF) experiments showed that both DROSHA and DGCR8 consistently colocalized with endogenous CGG nuclear RNA aggregates in brain sections of patients with FXTAS (Figures 2D and 2E; larger fields in Figures 2F and 2G), whereas DROSHA and DGCR8 were diffusely localized within the nucleoplasm of age-matched non-FXTAS controls. Finally, we tested the localization of endogenous DROSHA, DGCR8, and CGG aggregates in brain sections of a knockin mouse model, in which endogenous CGG repeats had been replaced with an expansion of 98 CGG repeats (Willemsen et al., 2003). RNA FISH coupled to IF labeling showed the presence of rare nuclear CGG RNA aggregates that colocalized with both endogenous DROSHA and DGCR8 in mice expressing expanded CGG repeats (Figures S2E and S2F). By contrast, DROSHA and DGCR8 were diffuse throughout the nucleoplasm in control mice. We noted that the CGG RNA aggregates were larger and much more frequent in patients with FXTAS than in knockin mice, which is consistent with the milder neurological disturbances observed in knockin mice compared to patients with FXTAS (Willemsen et al., 2003).

The direct binding of DGCR8 to expanded CGG trinucleotide repeats raises the question as to how specific this interaction is and, notably, whether DGCR8 can also recognize other trinucleotide repeats such as expanded CUG repeats, the mutation responsible of myotonic dystrophy of type 1 (DM1). UV-crosslinking assays showed that DGCR8 binds weakly to RNAs containing 20 or 100 CUG repeats, compared to an RNA containing 100 CGG repeats (Figure S2G). Similarly, both DROSHA and DGCR8 from mouse brain extract were captured by RNA affinity columns composed of expanded CGG repeats, yet neither was recruited by expanded CUG repeats (Figure S2H). Finally, RNA FISH coupled to IF demonstrated that both endogenous DROSHA and DGCR8 were recruited within CGG RNA repeat aggregates that formed in COS7 cells transfected with a plasmid expressing 60 CGG repeats (Figures S2I and S2J). In contrast, DROSHA and DGCR8 were not recruited within RNA aggregates of similarly transfected cells expressing either expanded CUG or AUUCU repeats, which are involved in DM1 and spinocerebellar ataxia of type 10 (SCA10), respectively (Figures S2I and S2J). Identical results were obtained in neuronal PC12 or GT17 cells (data not shown). These results indicate that DGCR8 binds preferentially expanded CGG repeats compared to CUG repeats, which is consistent with the enhanced stability of the double-stranded helical structure formed by expanded CGG repeats compared to CUG repeats (Mooers et al., 2005; Sobczak et al., 2003; Zumwalt et al., 2007; Kiliszek et al., 2009, 2011; Kumar et al., 2011). Alternatively, the presence of U:U mismatches versus noncanonical G:G base pairing, or other structural differences that are significant, between CUG and CGG hairpins, may impair the binding of DGCR8. Overall, these results suggest that DROSHA and DGCR8 are specific components of the CGG RNA aggregates in FXTAS.

DROSHA Does Not Cleave Expanded CGG Repeats

The binding of DROSHA and DGCR8 to the expanded CGG RNA repeat raises the possibility that DROSHA may cleave these repeats into shorter CGG hairpins, which is an attractive hypothesis considering that DICER was reported, in vitro, to partially process expanded CGG repeats into potentially toxic miRNA-like CGG RNA repeats (Handa et al., 2003), an observation also reported for expanded CUG and CAG repeats (Krol et al., 2007; Bañez-Coronel et al., 2012). Accordingly, we tested whether DROSHA and DGCR8 can process an RNA containing 60 expanded CGG repeats. However, no cleavage products were observed, although DROSHA correctly processed a control pri-miR-125 (Figure S3A). Furthermore, we found no trace of small miRNA-like CGG RNA after massive parallel sequencing of RNA extracted from cells transfected with a plasmid expressing 60 CGG repeats (data not shown). We propose that structural differences between pri-miRNAs and CGG expanded repeats, such as noncanonical G:G base pairing, impair the cleavage activity of DROSHA. Overall, the recruitment of DROSHA and DGCR8, without the processing of CGG repeats and the subsequent release of the proteins, suggests that the aggregates of CGG repeats may act as molecular sink, titrating DROSHA and DGCR8 away from their normal functions.

DROSHA and DGCR8 Are Partially Sequestered by Expanded CGG Repeats

To test a potential sequestration of DROSHA and DGCR8, we first analyzed the effect of expanded CGG repeats on the RNA-binding activity of DGCR8. Gel shift experiments demonstrated that addition of increasing amounts of unlabeled expanded CGG RNA repeat progressively competed with the binding of recombinant purified DGCR8 to radioactively labeled pri-miR-125 (Figure 3A). As a control, addition of unlabeled expanded CUG repeats had little or no effect. We confirmed these results by UV-crosslinking assays, which demonstrated that addition of increasing amounts of expanded CGG repeats competed with the binding of DGCR8 to radioactively labeled pri-miR-124 and pri-miR-125, whereas expanded CUG repeats had no effects (Figure S3B). Next, we tested the effect of

expanded CGG repeats on the processing activity of DROSHA. COS7 cells were cotransfected with a plasmid expressing an ectopic pri-miRNA under the expression of a CMV promoter, allowing the detection of the primary, precursor, and mature miRNAs by northern blotting (Figure 3B). Importantly, coexpression of increasing amounts of expanded CGG repeats reduced the processing of ectopic pri-miR-124 into pre-miR-124. This inhibition is specific because expression of control expanded CUG repeats had no effect on the biogenesis of pri-miR-124 (Figure 3B). We confirmed these results using ectopically expressed pri-miR-206, pri-miR-146, and pri-miR-26 and found that expression of expanded CGG repeats inhibited the DROSHA cleavage of pri-miRNA transcripts into pre-miRNAs (Figure S3C). As a control, expression of expanded CUG repeats had no or little effects on miRNA biogenesis (Figure S3D). These results suggest that expanded CGG repeats compete with the binding of DGCR8 to pri-miRNAs, reducing the quantity of free DROSHA and DGCR8 available to process pri-miRNAs, which leads to reduced processing of pri-miRNAs into pre-miRNAs.

Theoretically, the titration of free DROSHA and DGCR8 may result in reduced formation of mature miRNAs. To test this hypothesis, we quantified the expression of endogenous mature miRNAs upon expression of expanded CGG repeats in neuronal GT17 cells. Microarray profiling demonstrated that most of the miRNAs, which presented a modified expression, were reduced upon transfection of a plasmid expressing 60 CGG repeats (Figure 3C). Decreased levels of mature miRNAs were confirmed by quantitative real-time RT-PCR (Figure 3D). The limited number (56) of miRNAs presenting an expression change, as well as the limited decrease (~20%–50%) of miRNA levels, are consistent with a progressive titration of DROSHA and DGCR8 and with the early time point (24 hr after transfection) chosen for analysis. Also, we noted that the expression of a minority of miRNAs was upregulated upon expression of expanded CGG repeats. However, a similar upregulation occurred at early time point (24 hr after transfection) in neuronal GT17 or COS7 cells depleted of DGCR8 by siRNA, suggesting the existence of rescue mechanisms to transiently increase the expression of some miRNAs in response to DGCR8 reduction (Figures S3E and S3F).

Importantly, a decrease in miRNA expression could be triggered by a specific alteration of the miRNA-processing machinery but may also reflect a global alteration of RNA transcription due to reduced cell viability. To discriminate between these two hypotheses, we quantified the levels of the primary transcripts hosting the decreased miRNAs. Quantitative RT-PCR demonstrated that, whereas mature miRNAs presented reduced levels, expression of their corresponding pri-miRNAs was not altered or, even, increased (Figure 3E). Similarly, the expression of various mRNAs containing pri-miRNAs within their introns was not altered (Figure 3F). Comparable decrease of mature miRNA levels with no alterations of the expression of their corresponding pri-miRNAs was observed in a second cell model: COS7 cells expressing 60 CGG repeats (Figures S3G and S3H). These results demonstrate that expanded CGG repeats alter specifically the processing of pri-miRNAs, without affecting their transcription. As a further control, we tested the expression of mirtrons, which are miRNAs processed by the splicing machinery, thus independent of the processing by DROSHA (Okamura et al., 2007; Ruby et al., 2007). Quantitative RT-PCR analysis of mirtron-877, mirtron-1224, mirtron-1225, and mirtron-1226 levels in neuronal GT17 cells expressing expanded CGG repeats demonstrated that their expression was not altered (Figure 3G). Similarly, mirtron expression was not altered in COS7 cells expressing expanded CGG repeats (Figure S3I). Also, the mRNA and protein expression levels of DROSHA and DGCR8 were normal in cells expressing expanded CGG repeats, indicating that pathogenic CGG repeats affect the activity, but not the expression, of DROSHA and DGCR8 (Figures S3J and S3K). Finally, the transfection of a plasmid encoding DGCR8 in neuronal cells expressing expanded CGG repeats rescued the decreased expression of miRNAs, whereas expression of an inactive form of DGCR8, which was deleted for its double-stranded RNA-binding domains, presented no rescue activity (Figure 3H). Overall, these data suggest that expanded CGG repeats specifically reduced the activity of DROSHA and DGCR8, without affecting other mechanisms such as general transcription.

The Processing of miRNAs Is Altered in Brain Samples of Patients with FXTAS

The altered processing of miRNAs in cells overexpressing expanded CGG repeats raises the question whether a similar alteration occurs in patients with FXTAS. Microarray analysis of cerebellar samples from patients with FXTAS revealed that mis-regulation of miRNAs involved predominantly decreased expression compared to age-matched controls (Figure 4A). Quantitative RT-PCR analysis confirmed reduced quantities of various mature miRNAs in patients with FXTAS relative to controls (Figure 4B). Note that the expression of mature miRNAs was not fully abolished but decreased by 20%–50%, indicating a partial titration of DROSHA and DGCR8 by the expanded CGG repeats. In contrast, the quantity of mirtron-877, whose biogenesis depends on the splicing machinery and bypasses DROSHA, was normal in FXTAS brain samples (Figure 4B). As a control, quantification of the primary transcripts hosting the downregulated mature miRNAs demonstrated no changes or increased expression in patients with FXTAS (Figure 4C). We confirmed these results in frontal cortex samples of patients with FXTAS compared to age-matched controls. Quantitative RT-PCR confirmed that expression of mature miRNAs was decreased in patients with FXTAS (Figure 4D). Consistent with a specific alteration of the activity of DROSHA, the quantities of their corresponding pri-miRNAs were normal or increased in FXTAS samples compared to age-matched control samples (Figure 4E). As further controls, expressions of various mirtrons, whose biogenesis is independent of DROSHA, were normal in patients with FXTAS (Figure 4F). Overall, these results suggest that transcription and splicing of pri-miRNAs are not globally altered in patients with FXTAS but that their processing by DROSHA is reduced.

Overexpression of DGCR8 Rescues Neuronal Cell Death

A model of titration of DGCR8 by expanded CGG repeats in patients with FXTAS predicts that depletion of DGCR8 would enhance any phenotype caused by expanded CGG repeats. To test that hypothesis, we took advantage of

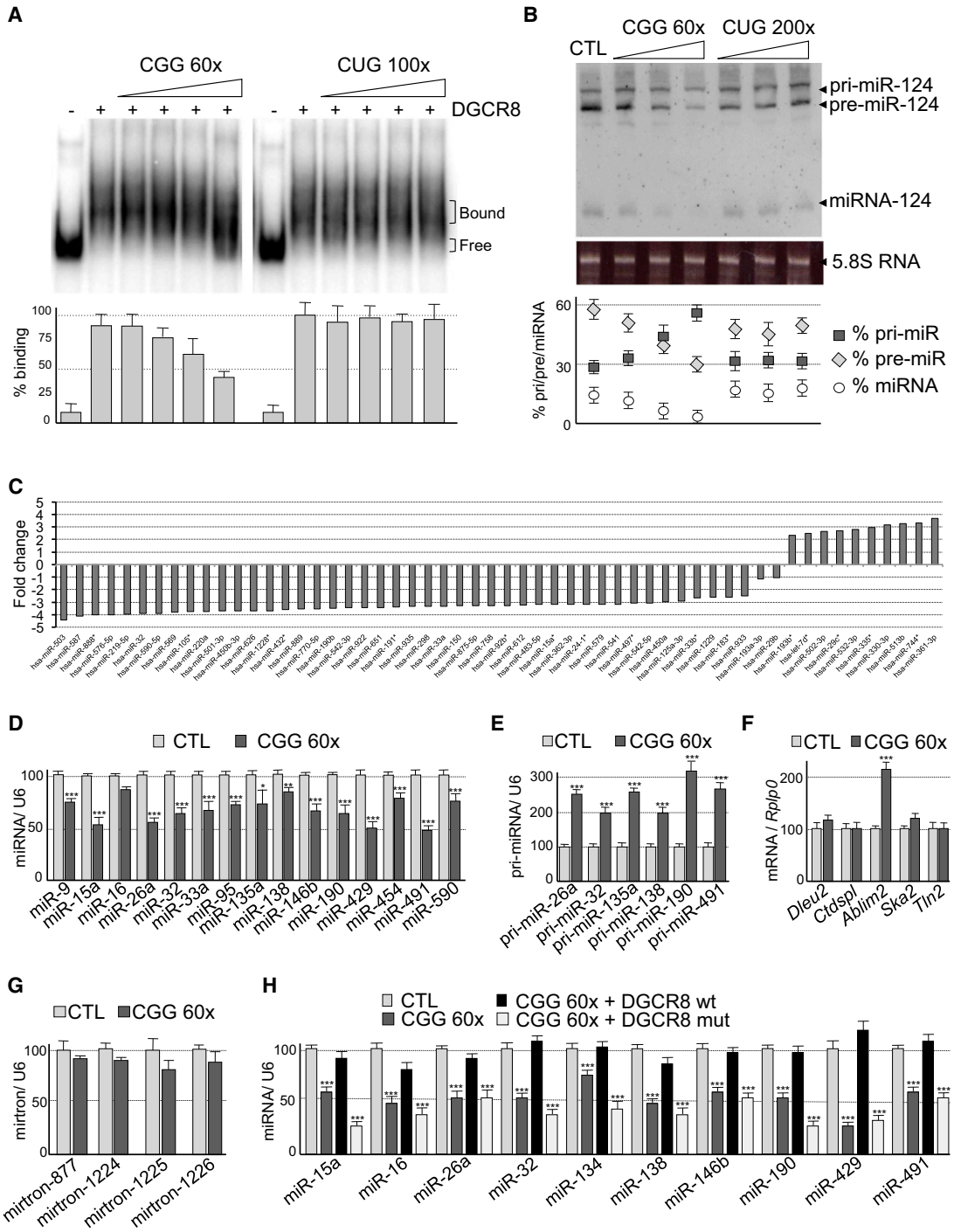


Figure 3. DROSHA and DGCR8 Activities Are Reduced in CGG-Expressing Cells

(A) Gel shift assays of recombinant His-DGCR8Δ binding with 100 nM (30,000 cpm) of uniformly α-[^{32}P]CTP internally labeled in-vitro-transcribed pri-miR-125 in presence of increasing amounts of nonlabeled in-vitro-transcribed RNA containing either 60 CGG repeats or 100 CUG repeats.

(B) A total of 20 μg of total RNA extracted from COS7 cells cotransfected with a pri-miR-124-2 minigene and increasing amounts of a plasmid expressing either 60 CGG or 200 CUG repeats were analyzed by northern blotting using a miR-124 γ-[^{32}P]ATP-labeled antisense probe. The mean of at least three independent transfections is depicted as the percentage of pri-, pre-, and mature miR-124-2. Error bars indicate SD.

(C) Microarray profiling of mature miRNAs expressed in GFP-positive FACS-isolated GT17 neuronal cells cotransfected with a plasmid expressing GFP and either a plasmid expressing no repeats (n = 3) or 60 CGG repeats (CGG, n = 3). Ordinate is in Log2 scale.

(D–G) Quantitative RT-PCR analysis of the expression of mature (D), pri-miRNAs (E), mRNAs (F), and mirtrons (G) relative to the U6 snRNA or to the *RPLPO* mRNA in GFP-positive FACS-isolated GT17 neuronal cells cotransfected with a plasmid expressing GFP and either a plasmid expressing no repeats (n = 3) or a plasmid

(legend continued on next page)

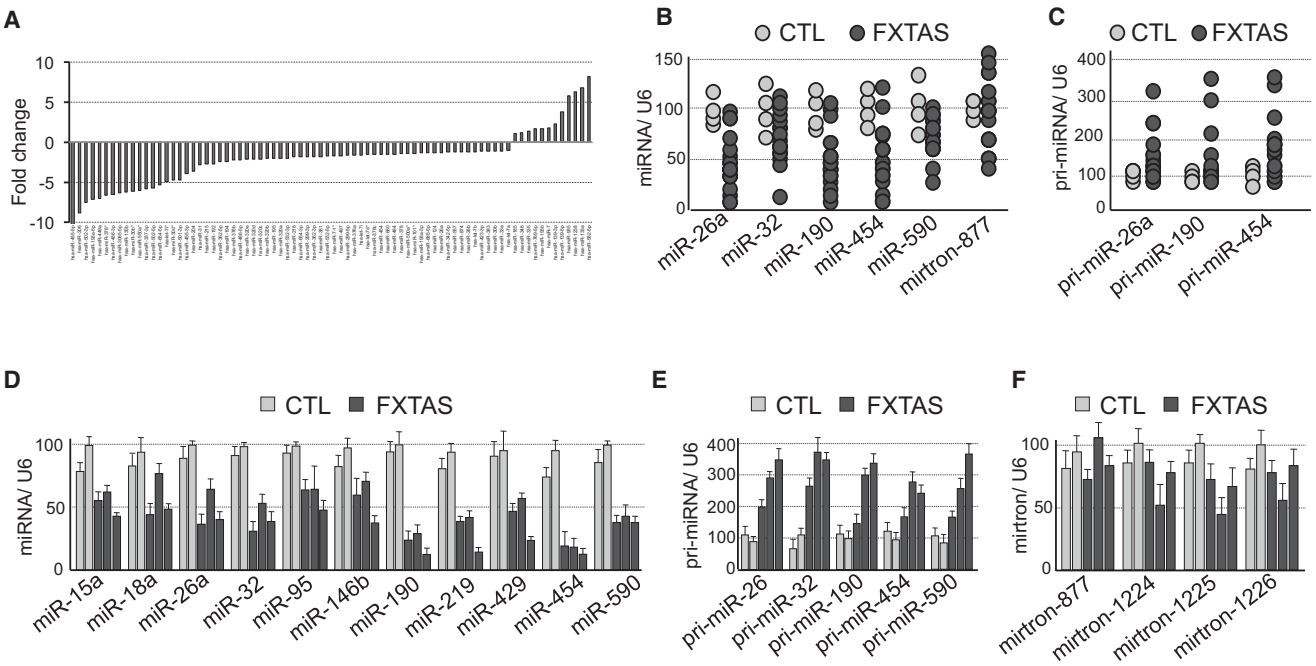


Figure 4. miRNA Levels Are Reduced in FXTAS Brain Samples

(A) Microarray profiling of mature miRNAs expressed in cerebellum samples of FXTAS and age-matched control patients.

(B and C) Quantitative RT-PCR analysis of the expression of mature (B) and pri-miRNAs (C) relative to the U6 snRNA in cerebellum samples of FXTAS (n = 13) and age-matched control (n = 4) patients.

(D–F) Quantitative RT-PCR analysis of the expression of mature (D), pri-miRNAs (E), and mirtrons (F) relative to the U6 snRNA in frontal cortex samples of FXTAS (n = 3) and age-matched control (n = 2) patients. Error bars indicate SD.

the neurodegenerative phenotype observed in transgenic *Drosophila melanogaster* expressing 90 CGG repeats (Jin et al., 2003). As previously described, expression of 90 CGG repeats decreases adult *Drosophila* viability (Figure S4A). Interestingly, decreased expression of Pasha (the *Drosophila* homolog of DGCR8) by RNAi in CGG transgenic flies resulted in enhanced early lethality (Figure S4A). A similar aggravated phenotype was observed with Drosha RNAi lines (Figure S4B). In contrast, overexpression of Drosha or Pasha in CGG flies had little or no effect and did not rescue fly lethality (data not shown). These data suggest that several pathological mechanisms may coexist in the fly model of FXTAS, a model consistent with the observation of cytoplasmic inclusions containing PURα in *Drosophila* (Jin et al., 2007), but not in human cells (Sellier et al., 2010). Accordingly, we tested mammalian neuronal cells, in which expression of expanded CGG repeats leads to formation of CGG RNA nuclear aggregates without formation of PURα-positive cytoplasmic aggregates. Organotypic cultures of E18 mouse cortex neurons were transfected at 7 days in vitro (DIV) with a plasmid expressing 60 CGG repeats and a plasmid expressing the GFP marker for 1 day (8 DIV). As reported previously by Chen et al. (2010), expression of expanded CGG repeats resulted in shorter dendrites and an ~40% decrease in dendritic branchpoints (Figures 5A and 5B). As a control, expression of expanded CUG repeats had no or little effect on neuronal dendritic complexity, indicating a specific deleterious effect of the expanded CGG repeats (Figure 5B). Importantly, the sole expression of a plasmid expressing DGCR8 restored normal dendritic growth and branching in neurons expressing the expanded CGG repeats (Figures 5A and 5B). In contrast, expression of an inactive form of DGCR8, which was deleted for its double-stranded RNA-binding domains, presented no rescue activity (Figure 5B). Similarly, expression of a control plasmid expressing MBNL1 had no significant effect and did not alleviate the dendritic alterations, indicating a specific role for DGCR8 (Figure 5B). We also tried to rescue the neuronal cell death induced by expression of expanded CGG RNA repeats by transfecting either miR-124, miR-9 or miR-125, which are important miRNAs for neuronal cell function, but we failed to observe a full rescue, suggesting that more than one miRNA is

expressing 60 CGG repeats (CGG, n = 3). *DLEU2* is the host gene of pri-miR-16-1, *CTDSPL* of pri-miR-26a1, *ABLIM2* of pri-miR-95, *SKA2* of pri-miR-454, and *TLN2* of pri-miR-190. Error bars indicate SD.

(H) Quantitative RT-PCR analysis of the expression of mature miRNAs relative to U6 snRNAs in GFP-positive FACS-isolated GT17 neuronal cells cotransfected with a plasmid expressing GFP, a plasmid expressing 60 CGG repeats, and either a vector expressing Flag-tagged wild-type (WT) DGCR8 or mutant (mut) DGCR8, deleted for its double-stranded RRM. Error bars indicate SD.

***p < 0.001; **p < 0.01; *p < 0.1.

See also Figure S3.

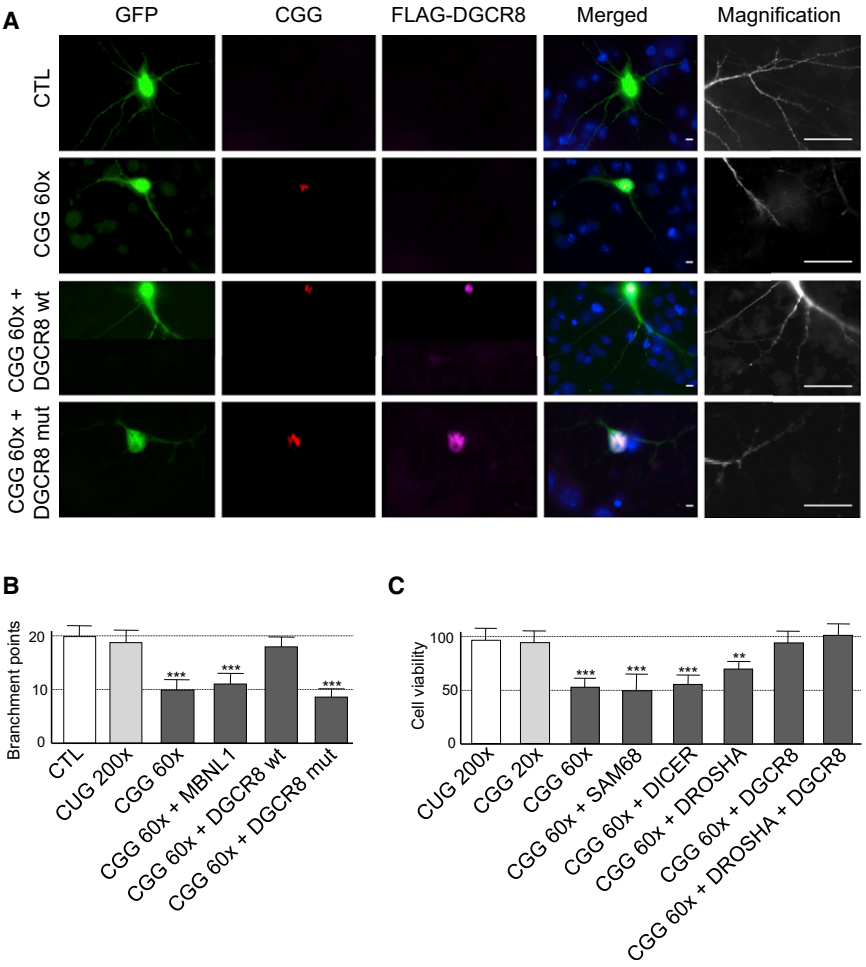


Figure 5. DGCR8 Is Sufficient to Rescue Expanded CGG Toxicity

(A) Primary cultures of cortex neurons from E18 mouse embryos were cotransfected with a plasmid expressing GFP, a vector expressing either no or 60 CGG repeats, and a plasmid expressing either a Flag-tagged wild-type DGCR8 or a mutant Flag-DGCR8 deleted for its double-stranded RRM. Neurons were analyzed 24 hr after transfection by RNA FISH using a CCG_{8x}-Cy3 DNA probe coupled to IF using an antibody directed against the FLAG tag. Magnification, 360×. Magnification of the insets, 2,500×. Scale bars, 20 μm.

(B) Quantification of the number of dendritic branchpoints of individual neurons transfected as in (A). Error bars indicate SD.

(C) Cell viability (tetrazolium) assay of neuronal GT17 cells transfected for 24 hr with a plasmid expressing either 200 CUG repeats or 20, 40, or 60 CGG repeats alone, or 60 CGG repeats plus a plasmid expressing either GFP-SAM68, Flag-DICER, Flag-DROSHA or Flag-DGCR8. Error bars indicate SD.

***$p < 0.001$; **$p < 0.01$.

See also Figure S4.

probably necessary to restore normal neuronal functions. Next, we tested whether DGCR8 can also rescue the neuronal cell death caused by expressing expanded CGG repeats. Transfection of a plasmid expressing 60 CGG repeats in neuronal GT17 cells resulted in a 50% decrease of the cell viability, whereas expression of either expanded CUG repeats or 20–40 control CGG repeats was not toxic (Figure 5C). Importantly, the overexpression of DGCR8 alone or of DGCR8 plus DROSHA alleviated the cell death due to the expanded CGG repeats (Figure 5C). As a control, transfection of control plasmids expressing either DROSHA alone, DICER or SAM68 had no or little effect. Similar results were obtained in COS7 cells (Figure S4C). Overall, these results suggest that a reduced quantity of free DROSHA and DGCR8 is the principal cause of the decreased dendritic complexity and reduced cell viability observed in neuronal cells expressing pathogenic-expanded CGG repeats.

DISCUSSION

Our results suggest a model for the cellular dysfunction induced by the expanded CGG repeats present within the 5′ UTR of the *FMR1* mRNA. Expanded CGG repeats form a double-stranded RNA hairpin (Sobczak et al., 2003; Zumwalt et al., 2007; Kumar et al., 2011; Kiliszek et al., 2011), which mimics the structure of pri-miRNAs recognized by DGCR8 (Zeng and Cullen, 2005; Han et al., 2006). We propose that DGCR8 interacts with the expanded pathogenic CGG repeats located within the 5′ UTR of the *FMR1* mRNA, thus sequestering itself and its partner, DROSHA (Figure 6). As a consequence, the level of free DROSHA-DGCR8 microprocessor is decreased, reducing the expression of mature miRNAs and ultimately resulting in neuronal cell dysfunction and degeneration. In that aspect, our work is reminiscent of the abnormal miRNA biogenesis and phenotypic abnormalities caused by the genetic haploinsufficiency of *Dgcr8* in the 22q11-deletion mouse model (Stark et al., 2008; Fénelon et al., 2011; Schofield et al., 2011).

A DROSHA-DGCR8 titration model has several predicted consequences for FXTAS pathology. First, a model based on DGCR8 titration predicts that the expansion of CGG repeats must exceed a minimal threshold size to accommodate DGCR8 binding (Zeng and Cullen, 2005; Han et al., 2006). This is consistent with the CGG repeat dependence of clinical involvement and degree of severity observed in patients with FXTAS (Tassone et al., 2007; Hoem et al., 2011), in whom it is estimated that "normal" CGG polymorphic repeat lengths are below 40 repeats, "gray zone" alleles contain 40–55 repeats, and patients with FXTAS are defined by premutation alleles containing 55–200 CGG repeats. Second, we observed that miRNA expression was decreased but not fully abolished, which is consistent with a partial and progressive sequestration of DROSHA-DGCR8 and, consequently, a progressive worsening of the disease symptoms. Third, previous studies have identified various

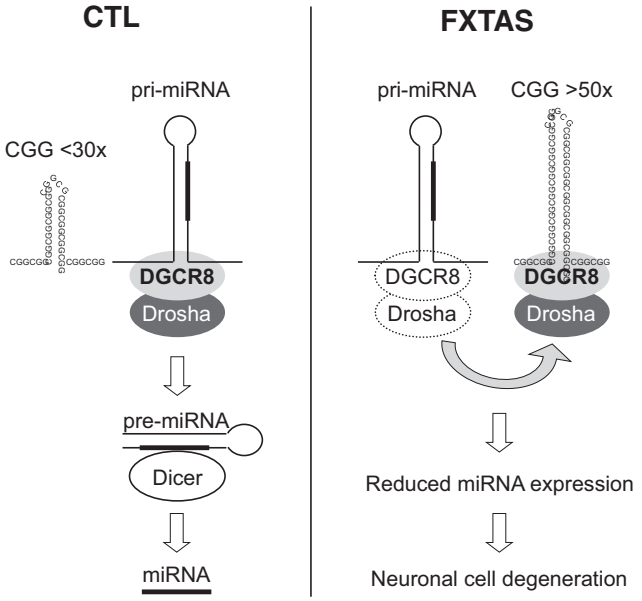


Figure 6. Model of DROSHA and DGCR8 Titration by Expanded CGG Repeats

Expanded CGG repeats fold into a double-stranded RNA hairpin that recruits DGCR8, resulting in the titration and immobilization of DROSHA and DGCR8. The reduced levels of free DROSHA and DGCR8 available to normally process pri-miRNAs into pre-miRNAs result in reduced levels of mature miRNAs, and ultimately, in neuronal cell dysfunctions.

proteins (e.g., PURα, hnRNP A2/B1, SAM68, etc.) as candidates for transduction of the CGG-expanded repeat toxicity (Iwahashi et al., 2006; Jin et al., 2007; Sofola et al., 2007; Sellier et al., 2010). However, the presence of cytoplasmic inclusions recruiting PURα in fly, but not in mammalian neuronal cells expressing expanded CGG repeats, raises the question of the superimposition of two different pathological mechanisms in *Drosophila*: one involving PURα in cytoplasmic inclusions, and another involving the sequestration of specific RNA-binding protein(s) by the expanded CGG repeats within nuclear aggregates. Such superimposition of pathogenic mechanisms may explain why overexpression of Pasha has no evident rescue effect in CGG-transgenic *Drosophila*, whereas expression of DGCR8 rescues the toxicity induced by expression of expanded CGG repeats in primary cultures of mouse neurons. Also, this superimposition model would explain the presence of ubiquitin-positive aggregates that do not colocalize with CGG RNA aggregates in knockin mouse models expressing expanded CGG repeats, as well as the recent report that only a subset of miRNAs is mis-regulated in *Drosophila* expressing expanded CGG repeats (Tan et al., 2012). Whether such a superimposition of pathological mechanisms also exists in patients with FXTAS remains an open question, yet to be determined. Finally, we found previously that SAM68 is sequestered within CGG RNA aggregates and that SAM68 rescues some of the splicing alterations observed in CGG-expressing cells (Sellier et al., 2010). However, we now show that the sequestration of SAM68 into CGG RNA aggregates requires DGCR8 and that restoration of SAM68 function is not sufficient to recover all normal neuronal cell functions. By contrast, expression of DGCR8 alone is sufficient to restore to normal both the dendritic morphological abnormalities and the loss of neuronal viability induced by expression of pathogenic-expanded CGG repeats in cultured mouse neurons.

In conclusion, this work raises the possibility that other human diseases could operate through a similar sequestration mechanism, provided that the mutant RNA forms a secondary structure that is recognized by the double-stranded RNA-binding protein, DGCR8. An appealing hypothesis in light of the recent finding that noncoding RNAs exceed coding transcripts by more than 5-fold and that an increasing number of potentially structured G-rich expanded repeats are found associated with pathologies, such as the expanded CCGGGG repeats in the *C9ORF72* gene that cause ALS-FTD (DeJesus-Hernandez et al., 2011). The current observations should also facilitate the development of novel model systems to better understand the molecular and cellular mechanisms underlying FXTAS. Finally, from the perspective of FXTAS treatment, identification of a key interaction between DGCR8 and the expanded CGG repeats represents an attractive target for therapeutic intervention. Indeed, if DROSHA and DGCR8 are sequestered, they are nevertheless present and potentially functional; hence, a strategy based on CGG-antisense oligonucleotides or pharmacological compounds (Disney et al., 2012) able to release the trapped DGCR8 would presumably return to normal the expression of miRNAs altered in FXTAS.

EXPERIMENTAL PROCEDURES

Nano-LC-MS/MS Analysis

Nuclear extract was prepared from mouse brain as described by Dignam et al. (1983). A total of 300 μg of nuclear extract was passed over an in-vitro-transcribed and -biotinylated RNA (Biotin 11 CTP; PerkinElmer) bound to streptavidin-coated magnetic beads (Dynabeads M-280 streptavidin; Invitrogen) in the presence of 20 mM HEPES, 300 mM NaCl, 8 mM $MgCl_2$, 0.01% NP40, 1 mM DTT, and protease inhibitor (PIC; Roche). The magnetic beads with immobilized RNA and its bound proteins were washed three times with the binding buffer, and bound proteins were eluted by boiling 3 min in the sample buffer prior to 4%–12% SDS-PAGE (NuPAGE 4%–12% bis-Tris Gel; Invitrogen) separation and silver staining (SilverQuest; Invitrogen). The protein bands were excised, digested, and identified using NanoESI_Ion Trap (LTQ XL; Thermo Fisher Scientific).

Cell Cultures and Transfections

Primary cortical neurons were prepared from C57Bl/6 mouse embryos at E18 and grown on polylysine-coated 24-well plates in neurobasal medium (NBM) supplemented with 1×B27, 0.5 mM L-glutamine, and 100 IU/ml penicillin/streptomycin at 37°C with 5% CO_2. Neurons were transfected at day 7 with Lipofectamine 2000 (Invitrogen) in 400 μl NBM. Medium was replaced after 3 hr with a 1:1 (v:v) mixture of conditioned and fresh NBM. After 30 hr, the neurons were fixed for FISH/ IF. COS7 cells were cultured in Dulbecco's modified Eagle's medium (DMEM), 10% fetal bovine serum, and gentamicin at 37°C in 5% CO_2. PC12 cells were cultured in DMEM, 10% horse serum, 5% fetal calf serum, and penicillin at 37°C, 5% CO_2. GT17 cells were grown in 10% fetal bovine serum, gentamicin, and penicillin at 37°C in 5% CO_2 and transfected 24 hr after plating in DMEM and 0.1% fetal bovine serum to block cell divisions, using either FuGENE HD (Roche) for COS7 cells or Lipofectamine 2000 for PC12 or GT17 cells. For RNA FISH/IF, GT17 cells were plated in 24-well plates on glass coverslips precoated with a solution of 1% collagen type I (BD Biosciences).

RNA FISH Combined with IF

Patients with FXTAS have been described previously (case 6, 7, and 9 of Greco et al., 2006). Mouse or human brain sections were deparaffinized

two times for 20 min in Histosol Plus (Shandon) and dehydrated as follows: twice in ethanol 100% (5 min), twice in ethanol 95% (5 min), once in ethanol 80% (5 min), once in ethanol 70% (5 min), and rinsed in PBS before RNA FISH. Glass coverslips containing plated cells or brain sections treated as described above were fixed in cold acetone during 20 min at −20°C and washed three times with PBS. The coverslips or slides were incubated for 10 min in PBS plus 0.5% Triton X-100 and washed three times with PBS before prehybridization in 40% DMSO, 40% formamide, 10% BSA (10 mg/ml), 2× SCC for 30 min. The coverslips or slides were hybridized for 2 hr in 40% formamide, 10% DMSO, 2× SCC, 2 mM vanadyl ribonucleoside, 60 μg/ml tRNA, and 30 μg/ml BSA plus 0.75 μg (CCG)8×-Cy3 DNA oligonucleotide probe (Sigma-Aldrich). The coverslips or slides were washed twice in 2× SCC/50% formamide and twice in 2× SCC. Following FISH, the coverslips or slides were washed twice successively in 2× SCC/50% formamide, in 2× SCC, and in PBS. The coverslips or slides were incubated 2 hr with primary antibody against Drosha (1:100 dilution, AB12286; Abcam) or DGCR8 (1:100 dilution, HPA019965; Sigma-Aldrich). Slides or coverslips were washed twice with PBS before incubation with a goat anti-rabbit secondary antibody conjugated with Alexa Fluor 488 (1:500 dilution; Thermo Fisher Scientific) for 60 min, incubated for 10 min in 2× SCC/DAPI (1:10,000 dilution), and rinsed twice in 2 × SSC before mounting in Pro-Long media (Molecular Probes). Slides were examined using a fluorescence microscope (Leica), and identical exposure or microscope setting was used for control or FXTAS brain section analyses.

Quantitative Real-Time PCR

Total RNA from cells or patient brains, the latter obtained under approved IRB protocols (University of California, Davis), was isolated by TriReagent (Molecular Research Center). cDNAs were generated using the miScript II RT Kit (QIAGEN) for quantification of miRNAs or the Transcriptor High Fidelity cDNA Synthesis Kit (Roche Diagnostics) for quantification of mRNAs. qPCR of miRNAs was realized using the miScript Primer Assay (QIAGEN) and miScript Sybr Green PCR Kit (QIAGEN) in a LightCycler 480 (Roche) with 15 min at 94°C followed by 50 cycles of 15 s at 94°C, 20 s at 55°C, and 20 s at 72°C. U6 snRNA was used as standard. qPCR of mRNAs was realized using the LightCycler 480 SYBR Green I Master (Roche) in a LightCycler 480 with 15 min at 94°C followed by 50 cycles of 15 s at 94°C, 20 s at 58°C, and 20 s at 72°C. The primers are listed in Table S1. *RPLPO* mRNA was used as standard, and data were analyzed using the LightCycler 480 analysis software (2ΔCt method).

For additional details, please see the Extended Experimental Procedures.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, four figures, and two tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2013.02.004.

LICENSING INFORMATION

ACKNOWLEDGMENTS

We thank Tom Cooper (Baylor College of Medicine, Houston) and Joelle Marie (CNRS, Gif-sur-Yvette, France) for the gift of the CUG-expressing plasmids, Karen Usdin (NIH, Bethesda, MD, USA) for the gift of the CGG 65x-expressing plasmid, Narry Kim (University of Seoul, Seoul, Korea) for the gift of the Flag-DROSHA and Flag-DGCR8 WT or mutant plasmids, Peng Jin (Emory School of Medicine, Atlanta), who provided the FXTAS model flies, Scott Pletcher (University of Michigan, Ann Arbor, MI, USA), who provided the Geneswitch flies, and Claudio Sette (University of Tor Vergata, Roma) for the gift of the GFP-SAM68 construct. This work was supported by ANR GENOPAT grant P007942 (to N.C.-B.), AFM and Jérôme Lejeune funding (to N.C.-B.), Collège de France (to C.S.), NIH grant K08NS069809 (to P.K.T.), NIH-NINDS grant NS062411 (to R.W.), Netherlands Brain Foundation (F2012(1)-101) (to R.W.), NIH grant 1R01GM079235-01A2 (to M.D.D.), and the NIH Roadmap Initiative grants DE019583 and AG032119 (to P.J.H.). Experiments were performed by C.S., F.F., R.T., F.H., T.T., F.R., and V.A. Samples and patient data were obtained from F.T., P.J.H., and R.W. Data were analyzed by C.S., C.T., A.P., M.D.D., P.K.T., H.M., R.W., N.C.-B., and P.J.H. The study was designed and coordinated by N.C.-B.

Received: June 26, 2012
Revised: November 30, 2012
Accepted: February 1, 2013
Published: March 7, 2013

REFERENCES

Arocena, D.G., Iwahashi, C.K., Won, N., Beilina, A., Ludwig, A.L., Tassone, F., Schwartz, P.H., and Hagerman, P.J. (2005). Induction of inclusion formation and disruption of lamin A/C structure by premutation CGG-repeat RNA in human cultured neural cells. Hum. Mol. Genet. *14*, 3661–3671.

Bañez-Coronel, M., Porta, S., Kagerbauer, B., Mateu-Huertas, E., Pantano, L., Ferrer, I., Guzmán, M., Estivill, X., and Martí, E. (2012). A pathogenic mechanism in Huntington's disease involves small CAG-repeated RNAs with neurotoxic activity. PLoS Genet. *8*, e1002481.

Chen, Y., Tassone, F., Berman, R.F., Hagerman, P.J., Hagerman, R.J., Willemsen, R., and Pessah, I.N. (2010). Murine hippocampal neurons expressing Fmr1 gene premutations show early developmental deficits and late degeneration. Hum. Mol. Genet. *19*, 196–208.

Davis, T.H., Cuellar, T.L., Koch, S.M., Barker, A.J., Harfe, B.D., McManus, M.T., and Ullian, E.M. (2008). Conditional loss of Dicer disrupts cellular and tissue morphogenesis in the cortex and hippocampus. J. Neurosci. *28*, 4322–4330.

DeJesus-Hernandez, M., Mackenzie, I.R., Boeve, B.F., Boxer, A.L., Baker, M., Rutherford, N.J., Nicholson, A.M., Finch, N.A., Flynn, H., Adamson, J., et al. (2011). Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. Neuron *72*, 245–256.

Denli, A.M., Tops, B.B., Plasterk, R.H., Ketting, R.F., and Hannon, G.J. (2004). Processing of primary microRNAs by the Microprocessor complex. Nature *432*, 231–235.

De Pietri Tonelli, D., Pulvers, J.N., Haffner, C., Murchison, E.P., Hannon, G.J., and Huttner, W.B. (2008). miRNAs are essential for survival and differentiation of newborn neurons but not for expansion of neural progenitors during early neurogenesis in the mouse embryonic neocortex. Development *135*, 3911–3921.

Dignam, J.D., Lebovitz, R.M., and Roeder, R.G. (1983). Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. *11*, 1475–1489.

Disney, M.D., Liu, B., Yang, W.Y., Sellier, C., Tran, T., Charlet-Berguerand, N., and Childs-Disney, J.L. (2012). A small molecule that targets r(CGG)(exp) and improves defects in fragile X-associated tremor ataxia syndrome. ACS Chem. Biol. *7*, 1711–1718.

Entezam, A., Biacsi, R., Orrison, B., Saha, T., Hoffman, G.E., Grabczyk, E., Nussbaum, R.L., and Usdin, K. (2007). Regional FMRP deficits and large repeat expansions into the full mutation range in a new Fragile X premutation mouse model. Gene *395*, 125–134.

Fénelon, K., Mukai, J., Xu, B., Hsu, P.K., Drew, L.J., Karayiorgou, M., Fischbach, G.D., Macdermott, A.B., and Gogos, J.A. (2011). Deficiency of Dgcr8, a gene disrupted by the 22q11.2 microdeletion, results in altered short-term plasticity in the prefrontal cortex. Proc. Natl. Acad. Sci. USA *108*, 4447–4452.

Greco, C.M., Hagerman, R.J., Tassone, F., Chudley, A.E., Del Bigio, M.R., Jacquemont, S., Leehey, M., and Hagerman, P.J. (2002). Neuronal intranuclear inclusions in a new cerebellar tremor/ataxia syndrome among fragile X carriers. Brain *125*, 1760–1771.

Greco, C.M., Berman, R.F., Martin, R.M., Tassone, F., Schwartz, P.H., Chang, A., Trapp, B.D., Iwahashi, C., Brunberg, J., Grigsby, J., et al. (2006). Neuropathology of fragile X-associated tremor/ataxia syndrome (FXTAS). Brain *129*, 243–255.

Gregory, R.I., Yan, K.P., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N., and Shiekhattar, R. (2004). The Microprocessor complex mediates the genesis of microRNAs. Nature *432*, 235–240.

Hagerman, R.J., Leehey, M., Heinrichs, W., Tassone, F., Wilson, R., Hills, J., Grigsby, J., Gage, B., and Hagerman, P.J. (2001). Intention tremor, parkinsonism, and generalized brain atrophy in male carriers of fragile X. Neurology *57*, 127–130.

Han, J., Lee, Y., Yeom, K.H., Kim, Y.K., Jin, H., and Kim, V.N. (2004). The Drosha-DGCR8 complex in primary microRNA processing. Genes Dev. *18*, 3016–3027.

Han, J., Lee, Y., Yeom, K.H., Nam, J.W., Heo, I., Rhee, J.K., Sohn, S.Y., Cho, Y., Zhang, B.T., and Kim, V.N. (2006). Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. Cell *125*, 887–901.

Handa, V., Saha, T., and Usdin, K. (2003). The fragile X syndrome repeats form RNA hairpins that do not activate the interferon-inducible protein kinase, PKR, but are cut by Dicer. Nucleic Acids Res. *31*, 6243–6248.

Haramati, S., Chapnik, E., Sztainberg, Y., Eilam, R., Zwang, R., Gershoni, N., McGlinn, E., Heiser, P.W., Wills, A.M., Wirguin, I., et al. (2010). miRNA malfunction causes spinal motor neuron disease. Proc. Natl. Acad. Sci. USA *107*, 13111–13116.

Hashem, V., Galloway, J.N., Mori, M., Willemsen, R., Oostra, B.A., Paylor, R., and Nelson, D.L. (2009). Ectopic expression of CGG containing mRNA is neurotoxic in mammals. Hum. Mol. Genet. *18*, 2443–2451.

Hébert, S.S., Papadopoulou, A.S., Smith, P., Galas, M.C., Planel, E., Silahtaroglu, A.N., Sergeant, N., Buée, L., and De Strooper, B. (2010). Genetic ablation of Dicer in adult forebrain neurons results in abnormal tau hyperphosphorylation and neurodegeneration. Hum. Mol. Genet. *19*, 3959–3969.

Hoem, G., Raske, C.R., Garcia-Arocena, D., Tassone, F., Sanchez, E., Ludwig, A.L., Iwahashi, C.K., Kumar, M., Yang, J.E., and Hagerman, P.J. (2011). CGG-repeat length threshold for FMR1 RNA pathogenesis in a cellular model for FXTAS. Hum. Mol. Genet. *20*, 2161–2170.

Huang, T., Liu, Y., Huang, M., Zhao, X., and Cheng, L. (2010). Wnt1-cre-mediated conditional loss of Dicer results in malformation of the midbrain and cerebellum and failure of neural crest and dopaminergic differentiation in mice. J. Mol. Cell. Biol. *2*, 152–163.

Iwahashi, C.K., Yasui, D.H., An, H.J., Greco, C.M., Tassone, F., Nannen, K., Babineau, B., Lebrilla, C.B., Hagerman, R.J., and Hagerman, P.J. (2006). Protein composition of the intranuclear inclusions of FXTAS. Brain *129*, 256–271.

Jacquemont, S., Hagerman, R.J., Leehey, M., Grigsby, J., Zhang, L., Brunberg, J.A., Greco, C., Des Portes, V., Jardini, T., Levine, R., et al. (2003). Fragile X premutation tremor/ataxia syndrome: molecular, clinical, and neuroimaging correlates. Am. J. Hum. Genet. *72*, 869–878.

Jin, P., Zarnescu, D.C., Zhang, F., Pearson, C.E., Lucchesi, J.C., Moses, K., and Warren, S.T. (2003). RNA-mediated neurodegeneration caused by the fragile X premutation rCGG repeats in *Drosophila*. Neuron *39*, 739–747.

Jin, P., Duan, R., Qurashi, A., Qin, Y., Tian, D., Rosser, T.C., Liu, H., Feng, Y., and Warren, S.T. (2007). Pur alpha binds to rCGG repeats and modulates repeat-mediated neurodegeneration in a *Drosophila* model of fragile X tremor/ataxia syndrome. Neuron *55*, 556–564.

Kiliszek, A., Kierzek, R., Krzyzosiak, W.J., and Rypniewski, W. (2009). Structural insights into CUG repeats containing the 'stretched U-U wobble': implications for myotonic dystrophy. Nucleic Acids Res. *37*, 4149–4156.

Kiliszek, A., Kierzek, R., Krzyzosiak, W.J., and Rypniewski, W. (2011). Crystal structures of CGG RNA repeats with implications for fragile X-associated tremor ataxia syndrome. Nucleic Acids Res. *39*, 7308–7315.

Krol, J., Fiszer, A., Mykowska, A., Sobczak, K., de Mezer, M., and Krzyzosiak, W.J. (2007). Ribonuclease dicer cleaves triplet repeat hairpins into shorter repeats that silence specific targets. Mol. Cell *25*, 575–586.

Kumar, A., Fang, P., Park, H., Guo, M., Nettles, K.W., and Disney, M.D. (2011). A crystal structure of a model of the repeating r(CGG) transcript found in fragile X syndrome. ChemBioChem *12*, 2140–2142.

Landthaler, M., Yalcin, A., and Tuschl, T. (2004). The human DiGeorge syndrome critical region gene 8 and Its *D. melanogaster* homolog are required for miRNA biogenesis. Curr. Biol. *14*, 2162–2167.

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Rådmark, O., Kim, S., and Kim, V.N. (2003). The nuclear RNase III Drosha initiates microRNA processing. Nature *425*, 415–419.

Mooers, B.H., Logue, J.S., and Berglund, J.A. (2005). The structural basis of myotonic dystrophy from the crystal structure of CUG repeats. Proc. Natl. Acad. Sci. USA *102*, 16626–16631.

Okamura, K., Hagen, J.W., Duan, H., Tyler, D.M., and Lai, E.C. (2007). The mirtron pathway generates microRNA-class regulatory RNAs in *Drosophila*. Cell *130*, 89–100.

Ruby, J.G., Jan, C.H., and Bartel, D.P. (2007). Intronic microRNA precursors that bypass Drosha processing. Nature *448*, 83–86.

Schaefer, A., O'Carroll, D., Tan, C.L., Hillman, D., Sugimori, M., Llinas, R., and Greengard, P. (2007). Cerebellar neurodegeneration in the absence of microRNAs. J. Exp. Med. *204*, 1553–1558.

Schofield, C.M., Hsu, R., Barker, A.J., Gertz, C.C., Blelloch, R., and Ullian, E.M. (2011). Monoallelic deletion of the microRNA biogenesis gene Dgcr8 produces deficits in the development of excitatory synaptic transmission in the prefrontal cortex. Neural Dev. *6*, 11.

Sellier, C., Rau, F., Liu, Y., Tassone, F., Hukema, R.K., Gattoni, R., Schneider, A., Richard, S., Willemsen, R., Elliott, D.J., et al. (2010). Sam68 sequestration and partial loss of function are associated with splicing alterations in FXTAS patients. EMBO J. *29*, 1248–1261.

Sobczak, K., de Mezer, M., Michlewski, G., Krol, J., and Krzyzosiak, W.J. (2003). RNA structure of trinucleotide repeats associated with human neurological diseases. Nucleic Acids Res. *31*, 5469–5482.

Sofola, O.A., Jin, P., Qin, Y., Duan, R., Liu, H., de Haro, M., Nelson, D.L., and Botas, J. (2007). RNA-binding proteins hnRNP A2/B1 and CUGBP1 suppress fragile X CGG premutation repeat-induced neurodegeneration in a *Drosophila* model of FXTAS. Neuron *55*, 565–571.

Stark, K.L., Xu, B., Bagchi, A., Lai, W.S., Liu, H., Hsu, R., Wan, X., Pavlidis, P., Mills, A.A., Karayiorgou, M., and Gogos, J.A. (2008). Altered brain microRNA biogenesis contributes to phenotypic deficits in a 22q11-deletion mouse model. Nat. Genet. *40*, 751–760.

Tan, H., Poidevin, M., Li, H., Chen, D., and Jin, P. (2012). MicroRNA-277 modulates the neurodegeneration caused by Fragile X premutation rCGG repeats. PLoS Genet. *8*, e1002681.

Tassone, F., Hagerman, R.J., Loesch, D.Z., Lachiewicz, A., Taylor, A.K., and Hagerman, P.J. (2000). Fragile X males with unmethylated, full mutation trinucleotide repeat expansions have elevated levels of FMR1 messenger RNA. Am. J. Med. Genet. *94*, 232–236.

Tassone, F., Iwahashi, C., and Hagerman, P.J. (2004). FMR1 RNA within the intranuclear inclusions of fragile X-associated tremor/ataxia syndrome (FXTAS). RNA Biol. *1*, 103–105.

Tassone, F., Adams, J., Berry-Kravis, E.M., Cohen, S.S., Brusco, A., Leehey, M.A., Li, L., Hagerman, R.J., and Hagerman, P.J. (2007). CGG repeat length correlates with age of onset of motor signs of the fragile X-associated

tremor/ataxia syndrome (FXTAS). Am. J. Med. Genet. B Neuropsychiatr. Genet. *144B*, 566–569.

Wang, Y., Medvid, R., Melton, C., Jaenisch, R., and Blelloch, R. (2007). DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. Nat. Genet. *39*, 380–385.

Willemsen, R., Hoogeveen-Westerveld, M., Reis, S., Holstege, J., Severijnen, L.A., Nieuwenhuizen, I.M., Schrier, M., van Unen, L., Tassone, F., Hoogeveen, A.T., et al. (2003). The FMR1 CGG repeat mouse displays ubiquitin-positive intranuclear neuronal inclusions; implications for the cerebellar tremor/ataxia syndrome. Hum. Mol. Genet. *12*, 949–959.

Zeng, Y., and Cullen, B.R. (2005). Efficient processing of primary microRNA hairpins by Drosha requires flanking nonstructured RNA sequences. J. Biol. Chem. *280*, 27595–27603.

Zumwalt, M., Ludwig, A., Hagerman, P.J., and Dieckmann, T. (2007). Secondary structure and dynamics of the r(CGG) repeat in the mRNA of the fragile X mental retardation 1 (FMR1) gene. RNA Biol. *4*, 93–100.

Neuron
Article

CGG Repeat-Associated Translation Mediates Neurodegeneration in Fragile X Tremor Ataxia Syndrome

Peter K. Todd,[1,*] Seok Yoon Oh,[1] Amy Krans,[1] Fang He,[1] Chantal Sellier,[5,6] Michelle Frazer,[1] Abigail J. Renoux,[1,2] Kai-chun Chen,[1] K. Matthew Scaglione,[1] Venkatesha Basrur,[3] Kojo Elenitoba-Johnson,[3] Jean P. Vonsattel,[7] Elan D. Louis,[8] Michael A. Sutton,[2] J. Paul Taylor,[9] Ryan E. Mills,[4] Nicholas Charlet-Berguerand,[5] and Henry L. Paulson[1,*]
[1]Department of Neurology
[2]Department of Physiology
[3]Department of Pathology
[4]Departments of Human Genetics and Computational Medicine & Bioinformatics
University of Michigan, Ann Arbor, MI 48109 USA
[5]Department of Translational Medicine, IGBMC, Illkirch 67400, France
[6]College de France, Paris 75000, France
[7]Taub Institute and Department of Pathology
[8]Department of Neurology
Columbia University, New York, NY 10032, USA
[9]Developmental Neurobiology, St. Jude Children's Research Hospital, Memphis, TN 38105, USA
*Correspondence: petertod@umich.edu (P.K.T.), henryp@umich.edu (H.L.P.)
http://dx.doi.org/10.1016/j.neuron.2013.03.026

SUMMARY

Fragile X-associated tremor ataxia syndrome (FXTAS) results from a CGG repeat expansion in the 5′ UTR of *FMR1*. This repeat is thought to elicit toxicity as RNA, yet disease brains contain ubiquitin-positive neuronal inclusions, a pathologic hallmark of protein-mediated neurodegeneration. We explain this paradox by demonstrating that CGG repeats trigger repeat-associated non-AUG-initiated (RAN) translation of a cryptic polyglycine-containing protein, FMRpolyG. FMRpolyG accumulates in ubiquitin-positive inclusions in *Drosophila*, cell culture, mouse disease models, and FXTAS patient brains. CGG RAN translation occurs in at least two of three possible reading frames at repeat sizes ranging from normal (25) to pathogenic (90), but inclusion formation only occurs with expanded repeats. In *Drosophila*, CGG repeat toxicity is suppressed by eliminating RAN translation and enhanced by increased polyglycine protein production. These studies expand the growing list of nucleotide repeat disorders in which RAN translation occurs and provide evidence that RAN translation contributes to neurodegeneration.

INTRODUCTION

A diverse group of human neurological disorders result from nucleotide repeat expansions (Orr and Zoghbi, 2007). These mutations can cause disease by protein gain-of-function, protein loss-of-function, or RNA gain-of-function mechanisms. For dominantly inherited repeat expansion disorders, defining whether the gain-of-function toxicity is elicited as RNA or as protein has traditionally depended on whether the repeat resides in an open reading frame (ORF) within an exon. For example, in Huntington's disease and other polyglutamine neurodegenerative disorders, expansion of exonic CAG repeats encoding polyglutamine promotes aggregation and alterations in the native properties of disease proteins (Orr and Zoghbi, 2007). In contrast, in myotonic dystrophy type 1, a CUG repeat expansion in the 3′ UTR of the *DMPK* gene causes toxicity predominantly as RNA (Cooper et al., 2009). The CUG repeat forms a hairpin structure that binds and sequesters certain splicing factors while also triggering activation of other pathogenic cascades.

Recently, however, the line separating RNA and protein gain-of-function nucleotide repeat diseases has begun to blur. RNA-mediated toxicity has now been proposed to contribute to polyglutamine diseases (Li et al., 2008) and bidirectional transcription through expansions can lead to repeats in both "coding" and "noncoding" mRNAs, raising the possibility that multiple toxic species may be produced from a single expansion (Ladd et al., 2007; Moseley et al., 2006; Wilburn et al., 2011). Moreover, evidence now suggests that repeats can be translated into proteins even if they do not reside in an AUG-initiated open reading frame (Zu et al., 2011). This repeat-associated non-AUG-initiated (RAN) translation can occur in all three possible ORFs of a given transcript, leading to numerous potentially toxic entities from a given repeat sequence (Pearson, 2011). RAN translation was recently shown to occur through the C9orf72 GGGGCC repeat expansion that causes ALS and frontotemporal dementia (Ash et al., 2013; DeJesus-Hernandez et al., 2011; Mori et al., 2013; Renton et al., 2011). These new findings raise key questions about how RAN translation occurs and whether it contributes directly to neurodegeneration. As the expected

mechanisms of toxicity differ depending on whether the inciting agent is RNA or protein, defining the critical toxic species in each repeat expansion disorder is an important step toward therapeutic development.

To explore the respective roles of RNA and RAN translation in repeat-associated neurodegeneration, we investigated fragile X-associated tremor ataxia syndrome (FXTAS), a common inherited cause of gait disorder, dementia, and tremor (Jacquemont et al., 2004). FXTAS is caused by a modestly expanded CGG nucleotide repeat (55–200) in the 5′ UTR of the fragile X mental retardation gene, *FMR1*. Much larger expansions of the same repeat cause fragile X syndrome, the most common inherited form of mental retardation, by silencing *FMR1* transcription (Penagarikano et al., 2007). By contrast, in FXTAS patients and animal models, the moderately expanded CGG repeat is associated with elevated FMR1 mRNA expression, neurodegeneration, and intranuclear neuronal inclusions that contain the CGG repeat mRNA and various proteins (Greco et al., 2006; Tassone et al., 2004). Research to date has focused on how the repeat might trigger neurodegeneration through an RNA mechanism (Jin et al., 2007; Sellier et al., 2010; Sofola et al., 2007), but critical aspects of disease pathology are not explained by a purely RNA-mediated process. Notably, the inclusions in FXTAS brains differ from those seen in other RNA-mediated disorders: they are large, ubiquitinated aggregates containing chaperone proteins such as HSP70 and many other proteins that do not interact directly with CGG repeat mRNA (Greco et al., 2006; Iwahashi et al., 2006). The inclusions of FXTAS instead more closely resemble neuronal intranuclear inclusions seen in polyglutamine diseases and other protein-mediated neurodegenerative disorders (Williams and Paulson, 2008).

Here we explain this paradox. We demonstrate that the CGG repeat expansion in FXTAS triggers RAN translational initiation within the 5′ UTR of FMR1 mRNA through an AUG-independent mechanism. The translated product, a cryptic polyglycine-containing protein we name FMRpolyG, is toxic in *Drosophila* and in human cell lines, capable of driving intranuclear inclusion formation, and present in FXTAS patient brains. The ability to produce FMRpolyG also explains pathologic discrepancies between two mouse models of FXTAS and directly influences the toxicity of CGG repeat constructs in *Drosophila*. Our findings support a disease model in which RAN translation of an expanded polyglycine protein contributes to FXTAS disease pathogenesis and suggest novel approaches toward therapeutic development in this and other neurodegenerative disorders.

RESULTS

Repeat-Associated Non-AUG-Initiated Translation and Inclusion Formation in a *Drosophila* Model of FXTAS

To explore the mechanism of inclusion formation in FXTAS, we utilized a *Drosophila* model of CGG repeat-mediated neurodegeneration in which the 5′ UTR from an FXTAS patient containing 90 CGG repeats is placed upstream of the coding region for GFP (Figure 1A; Jin et al., 2003; Todd et al., 2010). Initially designed to evaluate RNA-mediated toxicity, the $(CGG)_{90}$ GFP-expressing flies exhibit repeat length-dependent retinal degeneration (Jin et al., 2003). Remarkably, GFP-positive inclusions accumulate in $(CGG)_{90}$ GFP-expressing flies but not in flies expressing GFP alone (Figure 1B). These inclusions form in both the nucleus and cytoplasm and immunostain positively for ubiquitin and the chaperone HSP70 (Figures 1C and 1D).

CGG repeat RNA forms foci in FXTAS patients and in cell models of disease (Sellier et al., 2010; Tassone et al., 2004). We therefore evaluated whether the observed GFP inclusions in $(CGG)_{90}$ GFP-expressing flies colocalize with RNA foci. Multiple nuclear and cytoplasmic RNA foci were observed in retinal sections probed with a Cy5-$(CCG)_6$ RNA probe (Figure 1E, see Figure S1A available online) (Sellier et al., 2010). Only a fraction (43%) of RNA foci colocalized with GFP-positive inclusions (Figure 1F).

In principle, the GFP inclusions could result from general impairment of the ubiquitin proteasome system (UPS) by CGG repeat-containing mRNA/protein complexes. Arguing against this possibility, however, is the fact that coexpression of the temperature-sensitive β2 proteasomal subunit mutant DTS7 with GFP did not result in GFP inclusions (Figure S1B and data not shown), whereas it did induce inclusion formation by a more aggregate-prone fluorescent reporter, DsRed (Figure S1C). We next crossed flies expressing $(CGG)_{90}$ GFP with flies expressing DsRed. In this cross, if GFP inclusions resulted from a general toxic effect of the repeat, then the aggregation-prone DsRed should also form inclusions that coaggregate with GFP. These flies, however, developed GFP-positive inclusions without coaggregation of DsRed, suggesting that inclusion formation requires the CGG repeat to be present in the same mRNA that encodes GFP (Figure S1D). The presence of GFP inclusions prompted us to test whether enhancing or suppressing protein quality control pathways could modulate retinal degeneration in $(CGG)_{90}$ GFP-expressing flies. Consistent with a protein-mediated effect, retinal degeneration was enhanced by co-expressing the temperature-sensitive proteasomal subunit mutation DTS5 (β6 subunit) and suppressed by coexpression of the chaperone protein HSP-70 (Figure 1G; Jin et al., 2003; data not shown).

A recent report suggests that CAG repeats can trigger unconventional translation initiation (RAN translation) in the absence of an AUG start codon (Zu et al., 2011). We therefore asked whether CGG repeats trigger RAN translation upstream of, and in frame with, the GFP coding sequence to generate a higher molecular weight (HMW) GFP fusion protein that is prone to aggregate. Indeed, western blot analysis of $(CGG)_{90}$ GFP *Drosophila* lysates revealed an additional GFP species ~12 kDa larger than GFP (Figure 1H, arrow). Sequence analysis ruled out an unexpected upstream ATG mutation in the GFP coding sequence or loss of the GFP stop codon as the basis for this HMW GFP (full sequence is in Figure S1E). Stringent immunoprecipitations from $(CGG)_{90}$ GFP lysates (Figure S1F) also excluded ubiquitination as the cause of the HMW GFP protein.

Tandem mass spectroscopy (MS) of GFP immunoprecipitates from $(CGG)_{90}$ GFP *Drosophila* lysates confirmed the presence of an unconventional translation product. Three peptides were identified that correspond to the predicted protein sequence downstream of the repeat if the FMR1 5′ UTR were translated (Figure 1I). These peptides were not detected in flies expressing

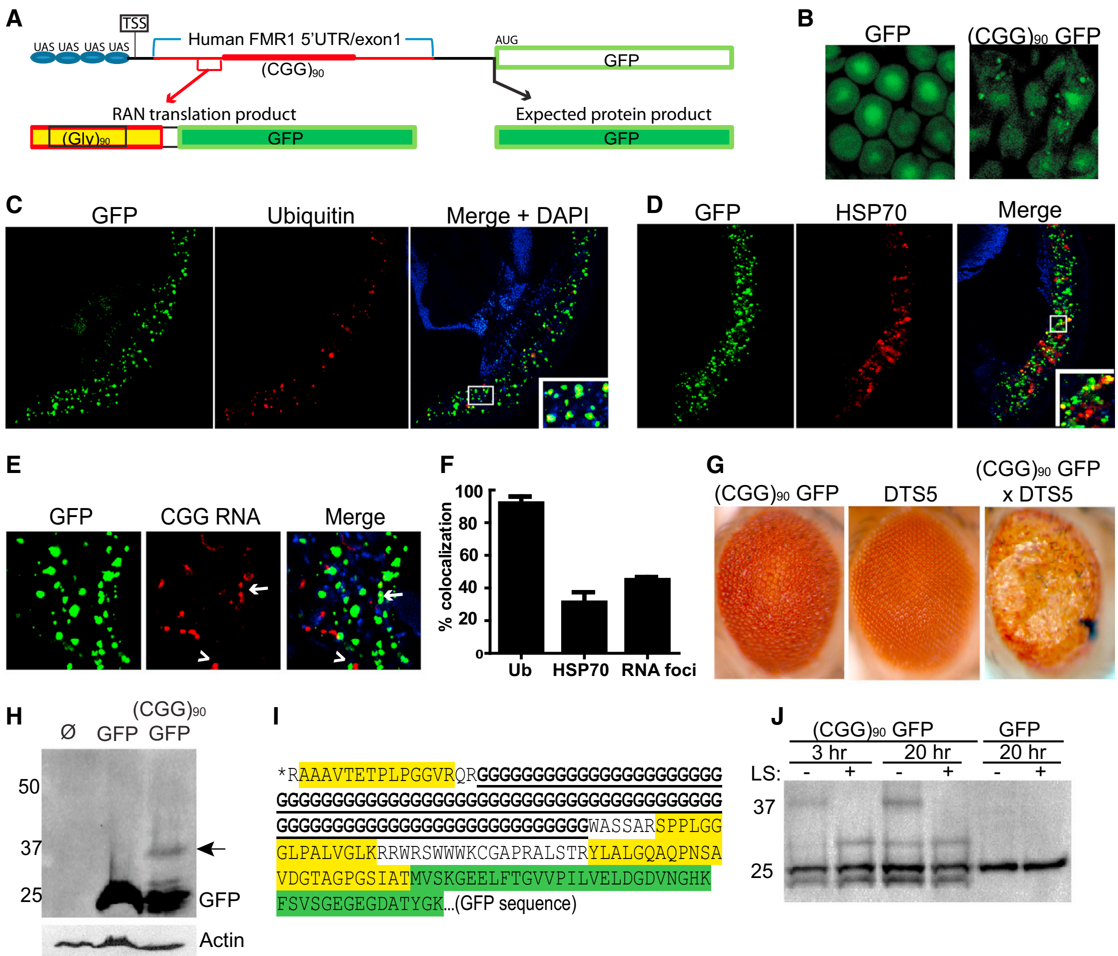


Figure 1. CGG RAN Translation in a *Drosophila* Model of FXTAS

(A) Schematic of $(CGG)_{90}$ GFP fly construct. A polyglycine protein is produced in these flies by RAN translation proceeding through the CGG repeat. Black sequence represents vector derived and red sequence represents human derived. Thicker red line indicates CGG repeat. The black arrow shows the expected AUG translational initiation site for GFP and the expected product. The red bracket and arrow show the RAN translational initiation region for the polyglycine-GFP fusion protein identified by tandem MS. TSS is the presumed transcription start site.

(B) GFP inclusions in oomatidia from $(CGG)_{90}$ GFP-, but not GFP-, expressing flies.

(C) Confocal micrographs of transverse retinal sections from gmr-GAL4; $(CGG)_{90}$ GFP flies reveal nuclear and cytoplasmic inclusions that colocalize with ubiquitin.

(D) $(CGG)_{90}$ GFP inclusions partially colocalize with HSP70.

(E) In situ hybridization using a $Cy5(CCG)_6$ RNA probe on transverse retinal sections. CGG RNA foci form in the nucleus and cytoplasm of $(CGG)_{90}$ GFP flies and are either distinct from (arrowhead), or overlap with (arrow), GFP inclusions.

(F) Quantification of colocalization of GFP aggregates with ubiquitin, HSP70, and CGG RNA foci. Error bars represent 95% confidence interval.

(G) Coexpression of proteasomal subunit mutant DTS5 with $(CGG)_{90}$ GFP enhances retinal degeneration at 28C.

(H) A HMW band is seen with anti-GFP antibody (arrow) in lysates from $(CGG)_{90}$ GFP-expressing flies. Lane 1, gmr-GAL4 flies (negative control); lane 2, gmr-GAL4; uas GFP; lane 3, gmr-GAL4; uas $(CGG)_{90}$ GFP.

(I) Tandem MS analysis of the HMW GFP band identifies three peptides (yellow) indicating that translation initiates above the repeat. Green sequence represents GFP. "*" indicates that predicted peptides above the indicated AA sequence were not detected.

(J) The HMW GFP product is selectively digested by the polyglycine endopeptidase, lysostaphin (LS).

GFP alone and are not predicted to exist in the *Drosophila* proteome. Based on the apparent molecular weight of the observed product, the identified peptide sequences corresponding to the 5′ UTR, and the reading frame of GFP, we conclude that the repeat is translated in the GGC reading frame to produce a 90 amino acid polyglycine stretch at the N terminus of the protein, with translation initiating just 5′ to the repeat. Consistent with this, further analysis identified a fourth peptide immediately N-terminal to the polyglycine repeat (Figure 1I) but no other peptides above this region.

No polyglycine fragment was detected by tandem MS, reflecting the lack of trypsin cleavage sites in expanded polyglycine. To confirm that the translation product contains polyglycine, we treated immunoprecipitates with lysostaphin, a specific

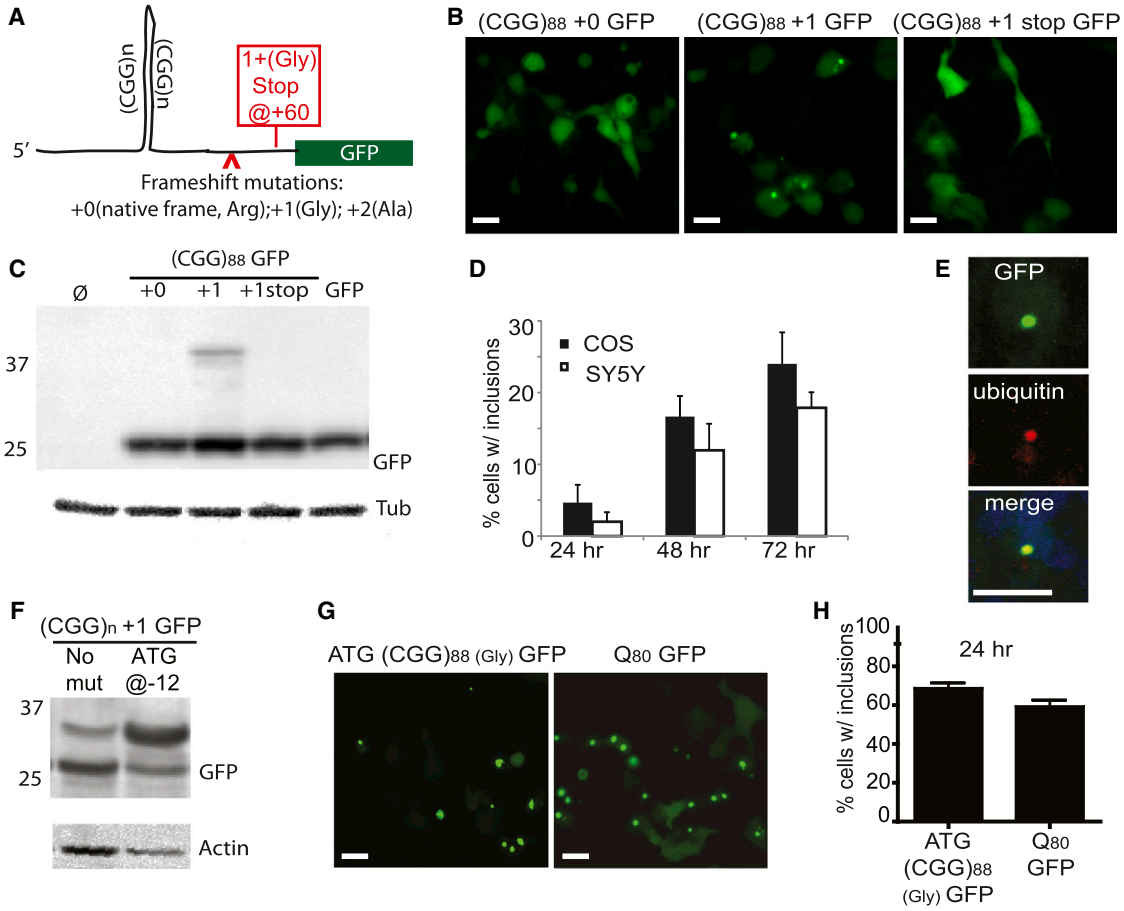


Figure 2. CGG Repeats Trigger RAN Translation and Inclusion Formation in Mammalian Cells

(A) Schematic of $(CGG)_{88}$ GFP vector and mutations introduced in various constructs. Arrowhead shows the site of additional base insertions to shift the frame of GFP relative to the repeat. Red box reflects stop codon introduced in the +1 (Gly) frame. Full sequences of all constructs are shown in Table S1.

(B) In COS cells 72 hr after transfection with $(CGG)_{88}$ GFP constructs, inclusions were observed when the CGG repeat was located in the +1(Gly) frame, but not in native CGG +0 (Arg) frame. Right panel includes a stop codon inserted between the repeat and the GFP coding sequence.

(C) COS cell lysates 72 hr after transfection with indicated plasmids, probed on western blot with antibodies to GFP or Tubulin.

(D) Quantification of inclusion formation by $(CGG)_{88}$ +1 GFP in COS or SY5Y neuroblastoma cells 24, 48, and 72 hr after transfection.

(E) Confocal microscopy showing inclusion in an SY5Y cell expressing $(CGG)_{88}$ +1 GFP, stained for ubiquitin (red) and costained for DAPI (blue).

(F) An ATG start site placed upstream of the repeat in the +1 (Gly) frame (ATG-$(CGG)_{88}$ +1 GFP) increases translation of the HMW GFP product.

(G) ATG-$(CGG)_{88}$ GFP enhances GFP inclusion formation similar to levels seen with a polyglutamine peptide fused to GFP (Q80-GFP), an aggregation-prone positive control.

(H) Comparison of percentage of GFP-positive cells with inclusions upon expression of ATG-$(CGG)_{88}$ GFP or Q_{80}-GFP 24 hr after transfection. For (D) and (H), error bars represent SEM.

pentaglycine endopeptidase (Huber and Schuhardt, 1970). Lysostaphin successfully cleaved the HMW GFP species from $(CGG)_{90}$ GFP lysates but had no effect on GFP alone, confirming the presence of a polyglycine repeat (Figure 1J).

RAN Translation Produces a Polyglycine-Containing Protein in Mammalian Cells

Zu et al. (2011) recently described RAN translation triggered by CAG/CUG repeat expansions in which translation initiates in all three possible reading frames, beginning within the hairpin itself. To test whether a similar phenomenon occurs with CGG repeat expansions, we transfected COS cells or SY5Y neuroblastoma cells with a construct containing the *FMR1* 5′ UTR from an FXTAS patient with 88 CGG repeats placed upstream of GFP (Figure 2A, Table S1). Importantly, this construct differs from the $(CGG)_{90}$ GFP fly sequence in that the GFP start codon resides in the CGG arginine-encoding frame relative to the repeat (specified as +0 in all figures). This frame recapitulates the relationship of the repeat to the FMRP ORF in FMR1 mRNA. In contrast, the GFP start codon in the *Drosophila* model resides in the GGC glycine-encoding frame (specified as +1 in all figures). For consistency, the repeat is referred to as CGG in all figures regardless of the frame in which it is translated, with modifiers placed before or after the repeat to indicate the relevant ORF and protein product and any introduced sequence changes. The sequences of all constructs are included in Table S1.

Expression of the $(CGG)_{88}$ +0 GFP construct led to diffuse GFP expression in transfected cells (Figure 2B), as previously reported (Arocena et al., 2005). To test whether RAN translation might initiate in other reading frames through the repeat, we added one or two bases between the CGG repeat and the ATG start codon of GFP: +1 (GGC, glycine-encoding) and +2 (GCG, alanine-encoding). As in $(CGG)_{90}$ GFP flies, the +1 (Gly) frame induced GFP inclusions in COS and SY5Y cells that accumulated over time (Figures 2B and 2D) and were predominantly intranuclear and ubiquitin positive (Figure 2E). Predictably, placing a stop codon after the repeat but before the GFP start site blocked GFP inclusion formation in the +1 (Gly) frame (Figure 2B). Consistent with the lack of an HMW FMRP species in FXTAS patients and animal models, no GFP-positive inclusions were identified in the +0 frame, despite the absence of intervening stop codons between the repeat and the ATG initiation codon of GFP. To test for the appearance of similar RAN translation products in mammalian cells as observed in *Drosophila*, we performed western blots on cell lysates 72 hr after transfection with each construct. An HMW GFP species was produced only in the +1 (Gly) frame and was no longer produced when a stop codon was placed between the repeat and the GFP start codon (Figure 2C). With the $(CGG)_{88}$ +1 GFP construct, this HMW GFP species constituted ~10% of the total cellular pool of GFP.

In a parallel set of experiments, we generated constructs in which GFP lacked its canonical start codon. As expected, this mutation markedly reduced the production of GFP in transfected cells, but when a 55 CGG repeat sequence was inserted upstream of this ATG-less GFP, placing the repeat in the glycine (+1)-encoding frame relative to GFP, the number of GFP-positive cells recovers to 60% of that seen with ATG-GFP (Figures S2A and S2B). In contrast, placing the CGG repeat in the arginine (+0)-encoding frame did not significantly increase the number of GFP-positive cells (Figures S2A and S2B). These findings correlated with production of RAN translation products as assessed by western blot: the HMW-GFP protein level was ~10% of that produced from GFP alone, while translation products from the +0 (Arg) construct were below the limit of detection (Figure S2C).

Studies comparing the aggregation properties of homopolymeric peptides suggest that small stretches of polyglycine are not prone to aggregation (Oma et al., 2004). To evaluate the aggregation properties of an expanded polyglycine-containing protein in FXTAS, we introduced an ATG start site upstream of the repeat in the glycine frame fused to GFP (ATG- $(CGG)_{88}$ +1 GFP). Incorporation of a canonical start site markedly increased production of the HMW GFP (Figure 2F), leading to inclusion formation in most cells 24 hr after transfection, which is comparable to the rate of inclusion formation seen when expanded polyglutamine is fused to GFP (Q_{80} GFP, Figures 2G and 2H).

Because the risk of developing FXTAS increases with larger repeats (Leehey et al., 2008), we evaluated the impact of repeat length on the production and aggregation of HMW GFP. The normal *FMR1* repeat in humans is between 20 and 45 CGGs, usually interrupted by one or two intervening AGGs. Constructs with 30, 50, or 88 CGG repeats in the +1 (Gly) frame with GFP all resulted in HMW GFP production (Figure 3A). Remarkably, production of the RAN translation product appeared to increase with decreasing repeat size, which may reflect differences in the transfer efficiency to PVDF membrane or greater translational efficiency for GFP and HMW GFP with shorter repeats (Chen et al., 2003). In contrast, aggregation decreased with decreasing repeat size, such that inclusions were infrequent at 50 repeats and nearly absent at 30 CGG repeats, suggesting a repeat length dependence to aggregation (Figures 3B and 3C). This repeat length dependence to inclusion formation was consistent across numerous cell types including primary cortical neurons (Figure S2D). Consistent with published studies, longer repeats were also associated with decreased cellular viability (Figure S2E; Arocena et al., 2005; Handa et al., 2005; Sellier et al., 2010).

RAN Translation of Polyglycine Protein Initiates 5′ to the CGG Repeat

To elucidate the mechanism by which RAN translation initiation occurs, we created a series of mutations in the sequence 5′ to the repeat to determine the minimal requirements for initiation. We first introduced a stop codon at −6 bp, −12 bp, −21 bp, or −63 bp from the start of the CGG repeat (Figure 3D). Stop codons placed at −6 bp or −12 bp prevented the appearance of GFP-positive inclusions and HMW GFP (Figures 3D–3F). In contrast, placement of a stop codon at −21 or −63 bp did not block production of HMW GFP, suggesting that RAN translation initiates between 21 and 12 bases 5′ of the CGG repeat (Figure 3D).

The bias toward translation in the glycine reading frame suggested preferred initiation at a specific non-AUG start site. A plausible explanation is the use of a specific alternative start codon as the translational origin rather than initiation within the hairpin itself. However, serially mutating each potential alternative start codon (i.e., "near-AUG" codons, differing by a single base from AUG) in the glycine frame within 60 bp 5′ of the CGG repeat did not eliminate production of HMW GFP (Figure 3G and data not shown). This result suggests either that a near-AUG codon is not needed to initiate translation or that multiple, different near-AUG codons proximal to the repeat can be utilized, so that eliminating any one near-AUG codon is not sufficient to prevent translation. To address this latter possibility, we placed a stop codon at −21 bp, which by itself does not prevent production of HMW GFP (Figure 3H), and mutated the only potential near-AUG codon downstream of this stop codon, a GUG codon at −11 bp. Mutating this GUG codon to GAG by itself had no impact on HMW GFP production but, combined with the stop codon at −21, HMW GFP production was lost (Figure 3H). This result suggests that, at least for some sequence contexts, a near-AUG codon close to the repeat is required for CGG RAN translation initiation, but the specific sequence 5′ to the repeat is less critical. To evaluate this, we deleted 48 nt just 5′ proximal to the CGG repeat, which impaired production of HMW GFP (Figure 3I). Deleting 91 nt 5′ proximal to the repeat nearly eliminated HMW GFP production, which could reflect the importance of the specific sequence just proximal to the repeat or represent a consequence of shortening the distance between the transcription start site and the repeat (Figure 3I).

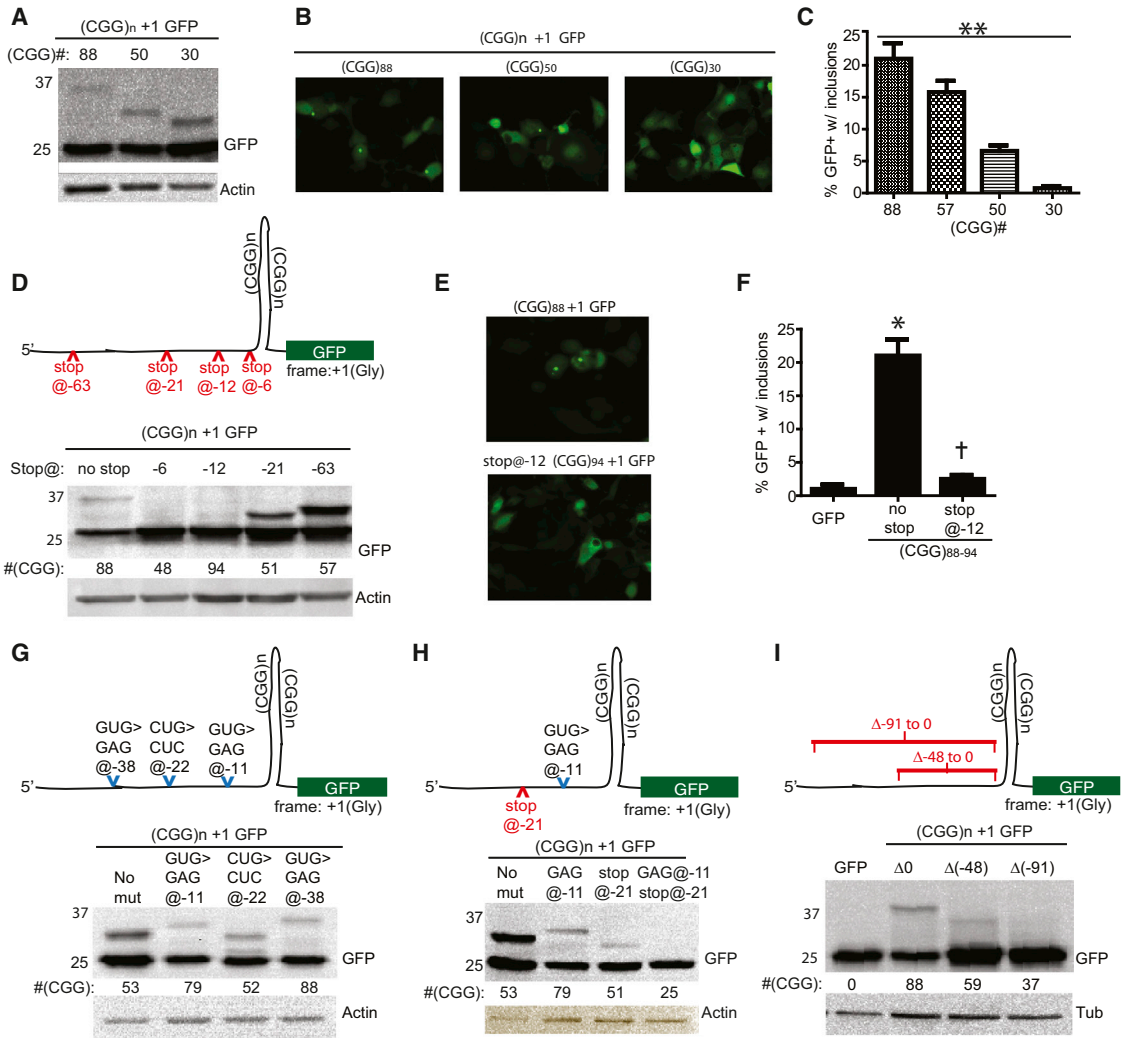


Figure 3. CGG RAN Translation in Glycine Reading Frame Occurs at Normal Repeat Lengths, Initiates before the Repeat, and Does Not Require a Specific Non-AUG Codon

(A) RAN translation occurs even with shorter repeats in the +1 (Gly) frame.

(B) Representative fluorescent micrographs of cells transfected with $(CGG)_n$ +1 GFP constructs with the indicated number of repeats.

(C) Percent transfected COS cells with GFP+ inclusions 72 hr after transfection of $(CGG)_n$ +1 GFP of the indicated repeat lengths.

(D) Top: schematic of $(CGG)_n$ +1 GFP construct with location of introduced stop codon mutations. Bottom: western blot of cell lysates 72 hr after transfection with the indicated constructs. Placing a stop codon 6 or 12 bp 5′ to the repeat inhibits production of HMW GFP.

(E) Fluorescent micrographs of COS cells expressing $(CGG)_{88}$ +1 GFP or stop@-12 $(CGG)_{94}$ +1 GFP, 72 hr after transfection.

(F) Quantification of GFP inclusion formation in the presence or absence of a stop codon 12 bp 5′ to the CGG repeat.

(G) Top: schematic demonstrating position of specific mutations in "near-AUG" codons that might serve as alternative start sites for CGG RAN translation in the +1 frame. Bottom: western blot of lysates from cells expressing the indicated constructs demonstrates that eliminating any single near AUG codon is insufficient to block RAN translation.

(H) Top: schematic demonstrating position of stop codon and near AUG codon mutation introduced into $(CGG)_n$ +1 GFP construct. Bottom: the elimination of a near AUG codon at −11 (lane 2) or the presence of a stop codon at −21 (lane 3) allows HMW GFP translation, but combining these mutations (lane 4) eliminates HMW GFP production.

(I) Top: schematic demonstrating deletion mutations that remove 48 or 91 nt just 5′ proximal to the repeat. Bottom: western blot demonstrating that removal of proximal sequence partially or completely impedes RAN translation in the +1 (Gly) frame. For this and other figures, differing sizes of the HMW GFP-positive bands reflect repeat instability incurred during cloning; repeat size is shown below the GFP blot for each lane. Positions of specific mutations are defined relative to the 5′ start of the repeat. Full sequences of all constructs are in Table S1. **$p < 0.001$ for trend, one-way ANOVA, *$p < 0.001$ versus GFP, †$p < 0.001$ versus $(CGG)_{88}$ +1 GFP, t test. For (C) and (F), error bars represent SEM.

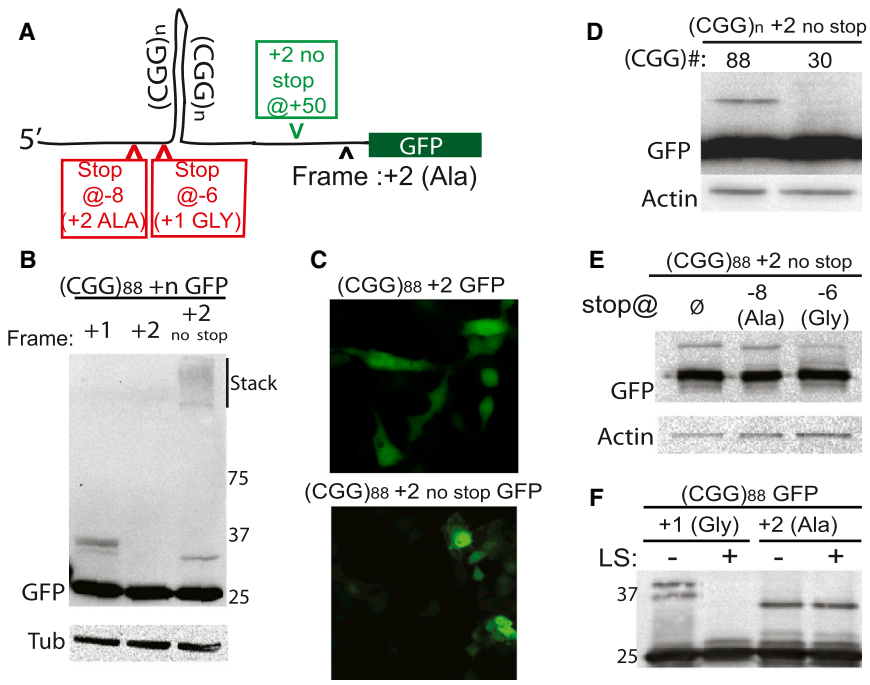


Figure 4. A RAN Translation Product Is Also Produced in the Alanine Frame

(A) Schematic of $(CGG)_{88}$ +2 GFP construct. Green arrowhead indicates where an intervening stop codon was removed. Red arrowheads indicate introduced stop codons.

(B) Unconventional translation resulting in a discrete HMW-GFP species and aggregated protein in the stack is detected in the +2 (GCG, alanine-encoding) frame when an intervening stop codon is removed.

(C) No GFP inclusions are observed in +2 (Ala) constructs even when the intervening stop codon is removed.

(D) HMW polyalanine product is absent with shorter CGG repeats.

(E) In +2 frame constructs, introduction of a stop codon at −8 bp (in the +2 frame) does not eliminate the RAN translation product. A stop codon at −6 bp (in the +1 frame) reduces but does not eliminate HMW GFP.

(F) Three hour digestion of GFP immunoprecipitates with Lysostaphin eliminates the HMW species in the +1 (Gly) frame but not in the +2 (Ala) frame.

RAN Translation Also Occurs in the Alanine Reading Frame

RAN translation associated with CGG repeats was not restricted to the glycine frame. In the +2 (GCG, alanine-encoding) frame, removing an in-frame stop codon between the repeat and the GFP coding sequence led to an HMW GFP band and aggregated protein (Figures 4A and 4B), although GFP-positive inclusions were not seen in transfected cells (Figure 4C). The +2 (Ala) frame HMW GFP species electrophoreses as a slightly smaller protein than what is seen with identical sized repeats in +1 (Gly) frame constructs, suggesting a different site of initiation (Figure 4B). In contrast to our results in the +1 (Gly) frame, reducing the repeat size to 30 CGGs eliminated expression of HMW GFP in the +2 (Ala) frame (Figure 4D). When a stop codon was introduced in the +2 (Ala) frame at −8 bp from an expanded $(CGG)_{88}$ repeat, expression of the HMW GFP persisted (Figure 4E). This result could reflect translational initiation in the +1 (Gly) frame followed by a frameshift into the +2 (Ala) frame, as frameshifts are known to occur with longer CAG repeat expansions (Stochmanski et al., 2012). To test this possibility, we introduced an upstream stop codon (−6 bp) in the +1 (Gly) frame while placing the downstream GFP sequence in the +2 (Ala) frame. This did not eliminate expression of an HMW GFP product (Figure 4E). To further exclude a Gly to Ala frame shift, we incubated lysates from cells expressing $(CGG)_{88}$ +1 GFP or $(CGG)_{88}$ +2 GFP with lysostaphin. Lysostaphin degraded HMW GFP produced in the +1 (Gly) frame but not in the +2 (Ala) frame (Figure 4F). We conclude that RAN translation associated with CGG repeats in the FMR1 5′ UTR occurs in at least two of three possible reading frames, but the constraints on translational initiation appear to differ for these two frames.

The FMR1 5′ UTR Is Engaged with Translating Ribosomes

CGG RAN translation in *Drosophila* and transfected mammalian cells suggests that unconventional translation may also occur in FXTAS patients. To explore this possibility, we first assessed whether the 5′ UTR of FMR1 mRNA is associated with translating ribosomes by querying ribosome profiling data sets previously generated in human cell lines (Guo et al., 2010; Hsieh et al., 2012; Ingolia et al., 2012). This technique combines ribosomal foot printing with next-generation sequencing to identify sites of active translation. In examining the distribution of ribosomes on FMR1 mRNA in published data sets, we observed that most sequence reads occurred, as expected, over coding regions of the FMR1 mature mRNA sequence (Figure 5A), with few if any reads in introns or the 3′ UTR. However, in the 5′ UTR of FMR1 mRNA, two peaks of protected sequence were present in the region just 5′ to the CGG repeat (Figure 5A), suggesting that fully assembled translating ribosomes do reside in this region of FMR1 RNA in human cells. These peaks exhibited ~40% of the mean read coverage and 60% of the peak read coverage observed in the first coding exon of FMR1 and were consistent across three published data sets in different human cell lines (Hsieh et al., 2012; Ingolia et al., 2012). Analysis of three mouse cell line data sets revealed a similar set of peaks just 5′ to the CGG repeat (Figures S4C and S4D) (Guo et al., 2010; Ingolia et al., 2011; Lee et al., 2012; Thoreen et al., 2012).

Data from studies utilizing the translational inhibitor Herringtonine to stall ribosomes at initiation suggest that many transcripts contain active upstream ORFs with initiation at alternative translational initiation sites (aTISs) (Figure S4) (Fritsch et al., 2012; Ingolia et al., 2011). Consistent with this observation, the average read density over the FMR1 5′ UTR was comparable

to the majority of transcripts across data sets in both human and mouse samples. We therefore focused our attention on the enhanced read density over near-AUG codons just 5′ proximal to the repeat, given that our cell culture data was most consistent with initiation in this region. We reasoned that enhanced read density could represent pausing of assembling ribosomes at initiation sites (Ingolia et al., 2011). The read density over two specific near-AUG codons (iGTG 12 nt 5′ to the repeat and iCTG 24 nt 5′ to the repeat) was significantly enhanced compared to other nucleotide triplets within the FMR1 5′ UTR (Figures S4A and S4B). Similarly, the mouse fmr1 5′ UTR also exhibited peaks of increased RP read density just 5′ proximal to the repeat that correlated with Herringtonine-identified initiation sites at near-AUG codons (Figures S4C and S4D). When compared to the distribution of read densities within 5′ UTRs on a transcriptome-wide level, these sites within the FMR1/fmr1 5′ UTR demonstrate relative read densities that are comparable to Herringtonine-confirmed alternative translational initiation sites (Figures S4E–S4G). Taken together, these data are consistent with translational initiation within the 5′ UTR of FMR1/fmr1 just proximal to the CGG repeat.

The FMR1 Polyglycine Protein Is Present in FXTAS Patient Brains

Translation of a 90 CGG repeat-containing FMR1 mRNA is predicted to produce an 11.5 kDa protein that contains an N-terminal polyglycine stretch followed by a 42 amino acid carboxyl terminal domain out of frame with the downstream FMRP start codon; we named this predicted protein FMRpolyG (Figure 5B). To determine whether FMRpolyG is made in FXTAS patients, we developed a monoclonal antibody (2J7) against a peptide from the predicted human protein (Figure 5B). Tested against recombinant FMRpolyG generated in bacteria as a GST-HIS fusion, 2J7 recognized two bands in bacterial lysates and purified protein samples (Figure 5C) that were confirmed by tandem MS to be FMRpolyG. In transfected mammalian cells expressing a FLAG-tagged FMRpolyG with 55 repeats, both 2J7 and FLAG antibody detected a protein electrophoresing at a slightly higher than expected MW of ~16 kDa (Figure 5D). In transfected COS cells, FLAG-FMRpolyG$_{55}$ displayed diffuse nucleocytoplasmic staining with occasional intranuclear inclusions detected by immunofluorescence with either 2J7 or anti-FLAG antibodies (Figure 5E). 2J7 immunostaining also colocalized with GFP inclusions formed in cells expressing FMRpolyG$_{100}$-GFP (Figure 5F). In contrast and as expected, 2J7 staining did not colocalize with GFP or inclusions in cells expressing an expanded polyglutamine-GFP fusion (Figure 5F).

To evaluate whether FMRpolyG is expressed in FXTAS patient brains, we performed western blots on cerebellar lysates from FXTAS patients. In pathologically confirmed FXTAS cases, an ~15 kDa band was identified in FXTAS lysates but not in control or AD brain lysates (Figure 5G, Figure S5E). We next evaluated whether this antibody differentially immunostained brain tissue from patients with clinically and pathologically confirmed FXTAS. Immunostaining with 2J7 was much more robust in FXTAS patient-derived hippocampal sections than in control tissue sections and included nuclear and perinuclear aggregates in FXTAS sections not seen in controls (Figure 5H, Figure S5A). In FXTAS hippocampus, numerous ubiquitin-positive inclusions were observed (Figure S6B), consistent with previous reports (Greco et al., 2006), and these inclusions coimmunostain with 2J7 (Figure 5I, Figures S5B–S5D). In contrast, 2J7 did not immunostain ubiquitinated polyglutamine inclusions in SCA 3 patient tissues (Figure 5I, Figure S6E). Similar staining by western blot and immunohistochemistry was observed using a different monoclonal antibody (2C13) against an overlapping epitope (Figures S5F–S5L).

We also generated an additional rabbit polyclonal antibody Ab605 raised against a larger peptide fragment of FMRpolyG (Figure 5B). This antibody also recognizes the recombinant protein (Figure 5J). To evaluate whether Ab605 recognizes FMRpolyG in inclusions in tissue, we first tested it on transverse retinal sections of *Drosophila* expressing the full antibody epitope. In flies expressing (CGG)$_{90}$ GFP, Ab605 readily colocalized with GFP+ inclusions (Figure S6A) but showed minimal staining in flies expressing GFP alone (Figure S6A). Whereas in control human tissue Ab605 displayed mild diffuse staining not seen in preimmune controls, Ab605 robustly stained neurons and intranuclear inclusions in FXTAS brain (Figure 5K, Figure S6C). By immunofluorescence, Ab605 staining colocalized with ubiquitin in FXTAS brain tissue (Figure S6D). Recognition of FXTAS inclusions was specific, as there was no staining by Ab605 of ubiquitinated polyglutamine inclusions in tissue sections from a spinocerebellar ataxia 3 brain (Figure S6E).

RAN Translation of the Polyglycine Protein Explains the Difference in Inclusion Formation between Two Mouse Models of FXTAS

To gauge the functional consequence of expressing a cryptic polyglycine protein, we turned to two similar mouse models of FXTAS. In both models, one generated in the Netherlands (Willemsen et al., 2003) and the other at the NIH (Entezam et al., 2007), premutation repeats were inserted into the 5′ UTR of the mouse *fmr1* locus. Both knockin (KI) models demonstrate intragenerational repeat instability and some evidence of neurodegeneration, but their phenotypes have not been directly compared (Brouwer et al., 2008). Comparing the cloning strategies used to make both lines, we noted that the NIH mouse model retains a greater region of mouse 5′ UTR surrounding the CGG repeat, including a TAA stop codon 18 bp 5′ of the repeat in the glycine frame (Figure 6A). This stop codon is not present in the Dutch knockin mouse or in humans. In cell culture experiments, placing the NIH mouse sequence, but not the Dutch mouse sequence, just proximal to the repeat blocked translation in the +1 (Gly) frame (Figure 6B). Thus, we would predict expression of the novel polyglycine protein only in the Dutch knockin mouse. Consistent with this prediction, 18-month-old mice from both lines differ greatly in the number and distribution of ubiquitinated inclusions. In Dutch knockin mice, ubiquitin-positive inclusions accumulate in the hypothalamus, cortex, and brainstem (Figures 6C and 6D) as previously reported (Brouwer et al., 2008), whereas they were seen less frequently in NIH knockin mice (Figures 6C and 6D). This difference exists despite similar expression of (CGG)$_n$ fmr1 RNA in both models (Figure 6E).

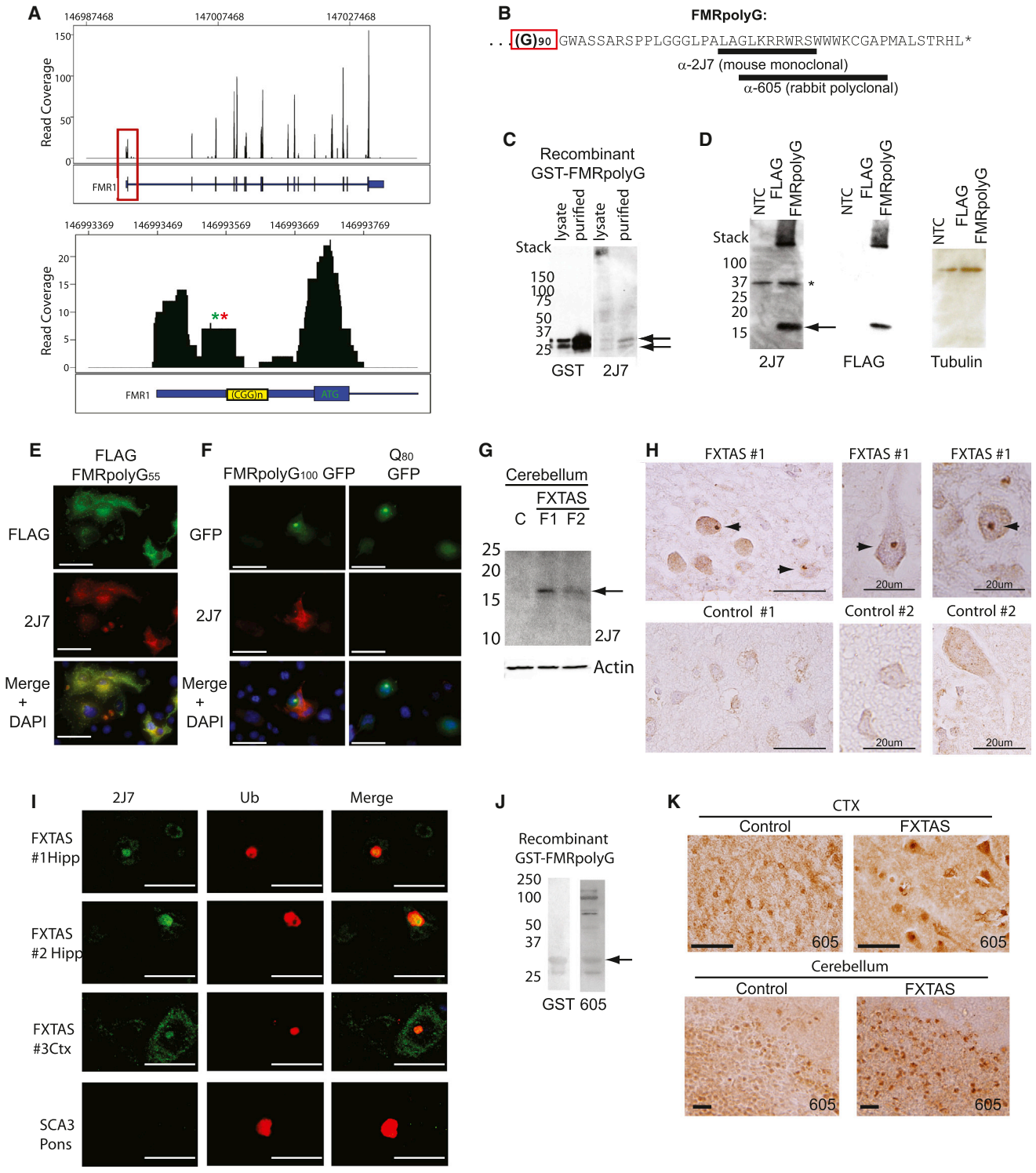


Figure 5. The Predicted Polyglycine Protein Is Present in FXTAS Patient Brains

(A) Read coverage map of *FMR1* locus derived from a published ribosomal profiling data set in HEK293 cells (Ingolia et al., 2012). Numbers along the x axis represent position within the genome. y axis represents number of sequence reads at each position. Black bars represent exons, intervening sequences are introns, and blue boxed sequences represent 5′ UTR (left) and 3′ UTR (right). Red box indicates the region shown at higher resolution in the bottom panel, which includes the FMR1 5′ UTR and first exon. Green and red asterisks indicate position of possible near AUG initiation codons iGUG (12 bp proximal to repeat) and iCTG (24 bp proximal to repeat), respectively. Note significant reads over region 5′ proximal to the CGG repeat sequence.

(legend continued on next page)

Because Ab605 was raised against a peptide sequence largely conserved in mouse, we used it to determine whether there is divergent immunostaining for FMRpolyG protein in the two models. Again, Dutch knockin mice show much greater immunostaining for FMRpolyG, including punctate nuclear staining consistent with inclusions in brain regions that also display ubiquitin-positive inclusions (Figures 6F and 6G). By coimmunofluorescence, Ab605 staining for FMRpolyG in Dutch knockin mice colocalized with ubiquitin-positive inclusions (Figure 6H). Together, these data suggest a dissociation of pathology in the two models based on differences in the ability to generate the polyglycine protein.

Translation of the Polyglycine Protein Contributes to CGG Repeat Toxicity in Human Cell Lines and *Drosophila*

The above results demonstrating RAN translation of a polyglycine protein in FXTAS models and patients support a role for FMRpolyG in aggregate formation in FXTAS. A critical, unanswered question is whether this polyglycine protein contributes to disease pathogenesis in FXTAS, or whether instead the CGG repeat as mRNA is wholly responsible for repeat-associated neurodegeneration. We therefore evaluated the effect of driving translation through the repeat on cell viability. Compared to expression of GFP alone, $(CGG)_{88}$ +1 GFP expression was associated with increased cell death at 72 hr, as measured by propidium iodide exclusion in GFP-positive cells (Figure 7A). This toxicity was repeat length dependent (Figure S2E), consistent with previous reports (Arocena et al., 2005; Sellier et al., 2010). When translation of the polyglycine protein was enhanced by placing an ATG upstream of the repeat, which increases production of the polyglycine protein without altering (CGG) GFP mRNA levels, the toxicity of the construct increased further (Figure 7A).

Because introducing a stop codon just before the repeat eliminated production of HMW GFP, we reasoned that this construct might allow us to determine whether initiating RAN translation is required for repeat-associated toxicity. Computer modeling of the FMR1 5′ UTR RNA secondary structure predicts an energetically favorable hairpin that includes the CGG repeat (Figure S3; Napierala et al., 2005). Placement of a stop codon at −12 bp relative to the repeat is not predicted to disrupt this hairpin (Figure S3), suggesting that CGG repeat structure should be preserved in this construct. We therefore measured cell death in transfected cells expressing $(CGG)_{88}$ +1 GFP or a similar $(CGG)_{94}$ +1 GFP construct containing a stop codon 12 bp before the repeat. Inclusion of this stop codon suppressed toxicity associated with the CGG repeat expansion, suggesting that a component of repeat toxicity reflects production of a polyglycine protein (Figure 7A).

To evaluate these effects in vivo, we generated a series of *Drosophila* lines in which the repeat was placed in different sequence contexts relative to GFP (Figure 7B, Table S2). We first regenerated lines in which the CGG repeat in the 5′ UTR of FMR1 is inserted upstream of GFP in the +1 (Gly) reading frame. In other lines, we inserted an ATG and FLAG tag just upstream of the CGG repeat to maximally drive expression of the polyglycine protein. To generate constructs in which the CGG repeat would be present as RNA but not translated into protein, we took two approaches: (1) inserting a stop codon 12 nt 5′ of the CGG repeat to prevent repeat-associated translation of the polyglycine protein, as shown earlier in cell culture (Figure 3D), or (2) moving the CGG repeat and surrounding regions of the 5′ UTR to a position downstream of GFP in the 3′ UTR. In cell culture, this repositioning blocked RAN translation (Figure 7C). All *Drosophila* lines expressed 100 CGG repeats, which were stable with intergenerational transmission.

Differential placement of the CGG repeat modestly altered transcript expression in *Drosophila*, with increased GFP mRNA in lines containing the repeat in the 5′ UTR (Figure S7A), perhaps due to local chromatin effects (Todd et al., 2010). Accordingly, we chose for further analysis lines in which GFP RNA production was comparable (Figure S7A). At the protein level, placing the CGG repeat in the 5′ UTR led to less overall GFP translation than did placement in the 3′ UTR (Figure 7D). As expected, including an ATG start codon upstream of the repeat led to increased production of HMW GFP but decreased production from the canonical TIS of GFP (Figure 7D).

To determine the impact of polyglycine protein expression on neurodegenerative phenotypes, we expressed each transgene in retinal oomatidia. In flies expressing the repeat in the 5′ UTR, we again observed the appearance of ubiquitin-positive GFP inclusions, which occurred more frequently when an ATG was placed upstream of the repeat (Figure 7E). In lines with a stop codon 5′ to the repeat or with the repeat positioned in the 3′ UTR, there were no GFP inclusions.

(B) Predicted sequence of human FMRpolyG protein with 90 glycines. Underlined regions represent peptides used to generate 2J7 and 605 antibodies. Polyglycine sequence is indicated by red box.
(C) GST and 2J7 antibody staining of recombinant purified GST-HIS-FMRpolyG$_{30}$ protein.
(D) Expression of a FLAG FMRpolyG$_{55}$ construct in COS cells stained with 2J7 and reprobed with anti-FLAG antibody and Tubulin. The protein runs higher than expected based on predicted size. NTC, no template control. *, nonspecific band. Arrow, FLAG-positive band.
(E) Immunofluorescence with FLAG (green) and 2J7 (red) in COS cells expressing FLAG-FMRpolyG$_{55}$.
(F) Coimmunofluorescence of GFP and 2J7 signal in COS cells expressing FMRpolyG$_{100}$ GFP (left panels) or Q$_{80}$ GFP (right panels).
(G) Western blot with 2J7 of cerebellar lysates from two FXTAS patients and an age-matched control. Arrow indicates bands seen in FXTAS but not control samples.
(H) Representative images of 2J7 immunostaining from frontal cortex (CTX) and hippocampus (Hipp) of control and FXTAS brain.
(I) Coimmunofluorescence with 2J7 (green) and anti-ubiquitin (red) in FXTAS hippocampus (top two panels) or cortex (third panel) from three different FXTAS subjects. In contrast, ubiquitinated inclusions in the pons of a patient with the polyglutamine disorder SCA-3 do not costain with 2J7.
(J) GST and rabbit polyclonal Ab605 staining of recombinant FMRpolyG protein. Arrow indicates band recognized by GST and Ab605.
(K) Representative images of Ab605 immunostaining of frontal cortex (CTX) and cerebellum from control and FXTAS brain. Unless otherwise noted, scale bars represent 50 μm.

We next determined whether expression of these different transgenes elicited a rough-eye phenotype as a measure of toxicity. Placing the repeat in the 5′ UTR and in-frame with GFP resulted in a moderate rough-eye phenotype (Figures 7F, 7G, and S8C), but this same repeat elicited only a very mild rough-eye phenotype when inserted into the 3′ UTR of GFP. Similarly, inserting a stop codon just 5′ to the repeat suppressed toxicity (Figures 7F, 7G, and S7C). In contrast, when an ATG was included 5′ to the repeat to drive polyglycine production, the rough-eye phenotype was more severe (Figures 7F and 7G). These results were consistent across multiple insertion lines (Figure S7C). When expressed ubiquitously, a CGG repeat placed in the 5′ UTR of GFP led to a decrease in viable progeny, and inclusion of an ATG 5′ of the repeat further enhanced this toxicity (Figure 7H). In contrast, including a stop codon 5′ of the repeat or placing the repeat in the 3′ UTR prevented CGG repeat-associated alterations in viability (Figure 7H). Together, these results suggest that RAN translation of a polyglycine protein contributes to CGG repeat toxicity in *Drosophila.*

DISCUSSION

Our results demonstrate that RAN translation occurs in association with CGG repeats in the neurodegenerative disorder FXTAS, a disease previously thought to result primarily from RNA-mediated toxicity. These findings, along with recent reports detailing unconventional translation through CAG repeats in spinocerebellar ataxia type 8 and myotonic dystrophy type I (Zu et al., 2011) and GGGGCC repeats in C9orf72-associated ALS/FTLD (Ash et al., 2013; Mori et al., 2013), suggest that RAN translation is a shared pathogenic mechanism in many repeat expansion disorders. We further demonstrate that production of one particular CGG RAN translation product in FXTAS, FMRpolyG, directly modulates CGG-associated pathology in two distinct model systems. First, the ability to generate the FMRpolyG protein explains a key pathologic discrepancy between two established knockin mouse models. Second, in *Drosophila* we demonstrate that CGG repeat-associated neurodegeneration is largely dependent on FMRpolyG production. These results suggest that RAN translation contributes to FXTAS pathogenesis (Figure 8) and support an emerging view that non-exonic repetitive elements can trigger toxicity simultaneously as both RNA and protein.

The mechanisms underlying RAN translation remain unclear. The unconventional translation described here appears to initiate predominantly at a near-AUG codon just 5′ proximal of the repeat. This finding suggests a model wherein a scanning 43S ribosomal preinitiation complex stalls at the CGG repeat, allowing for alternate usage of a near match at the initiation codon (Figure 8). This model is based on our observation that placing a stop codon just proximal to the repeat or shortening the 5′ leader before the repeat impairs RAN translation in this reading frame (Figure 3). In contrast, CGG RAN translation in the other two possible reading frames behaves differently. We do not detect any RAN translation product from the +0 (CGG, polyArginine) reading frame, and RAN translation in the +2 (GCG, polyAlanine) reading frame is less efficient, occurs when stop codons are inserted 5′ of the repeat, and demonstrates

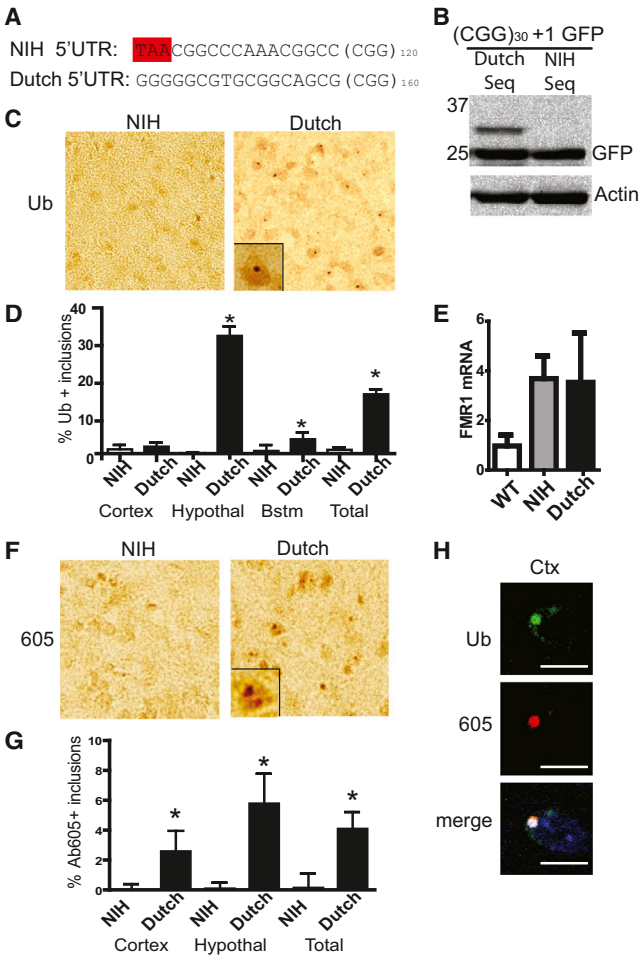


Figure 6. Sequence Differences 5′ of the Repeat Explain Divergent Inclusion Formation in Two Mouse Knockin Models of FXTAS

(A) Sequence differences in two established CGG knockin models of FXTAS highlighting a stop codon 18 bp before the repeat present only in the NIH mouse.

(B) Placement of the NIH mouse sequence, but not the Dutch sequence, just 5′ to the repeat eliminates the HMW GFP species in the +1 (Gly) frame.

(C) Representative images (original magnification 400×, inset 1,000×) of hypothalamus from 18-month-old NIH and Dutch knockin mice stained with antibody to ubiquitin.

(D) Quantification of ubiquitin-positive inclusions in 18-month-old NIH and Dutch mice in the specified brain regions.

(E) Relative expression of fmr1 mRNA in WT, NIH, and Dutch mice at 6 months of age (n = 4/genotype).

(F) Representative images of hypothalamus from 18-month-old NIH and Dutch knockin mice stained with Ab605 against FMRpolyG.

(G) Quantification of 605-positive inclusions in 18-month-old NIH and Dutch mice in the specified brain regions.

(H) Confocal microscopy in the Dutch knockin mice showing colocalization of ubiquitin and Ab605 staining inclusion in frontal cortex (left images). For (D), n > 300 cells/brain region and >1,000 cells/genotype, *p > 0.01 on Pearson's chi-square test. Error bars represent 95% confidence interval for (D) and (G) and SEM for (E).

CGG repeat length dependence (Figures 3, 4, and 5). Differences in the propensity for translational initiation in different reading frames were also reported for RAN translation of expanded

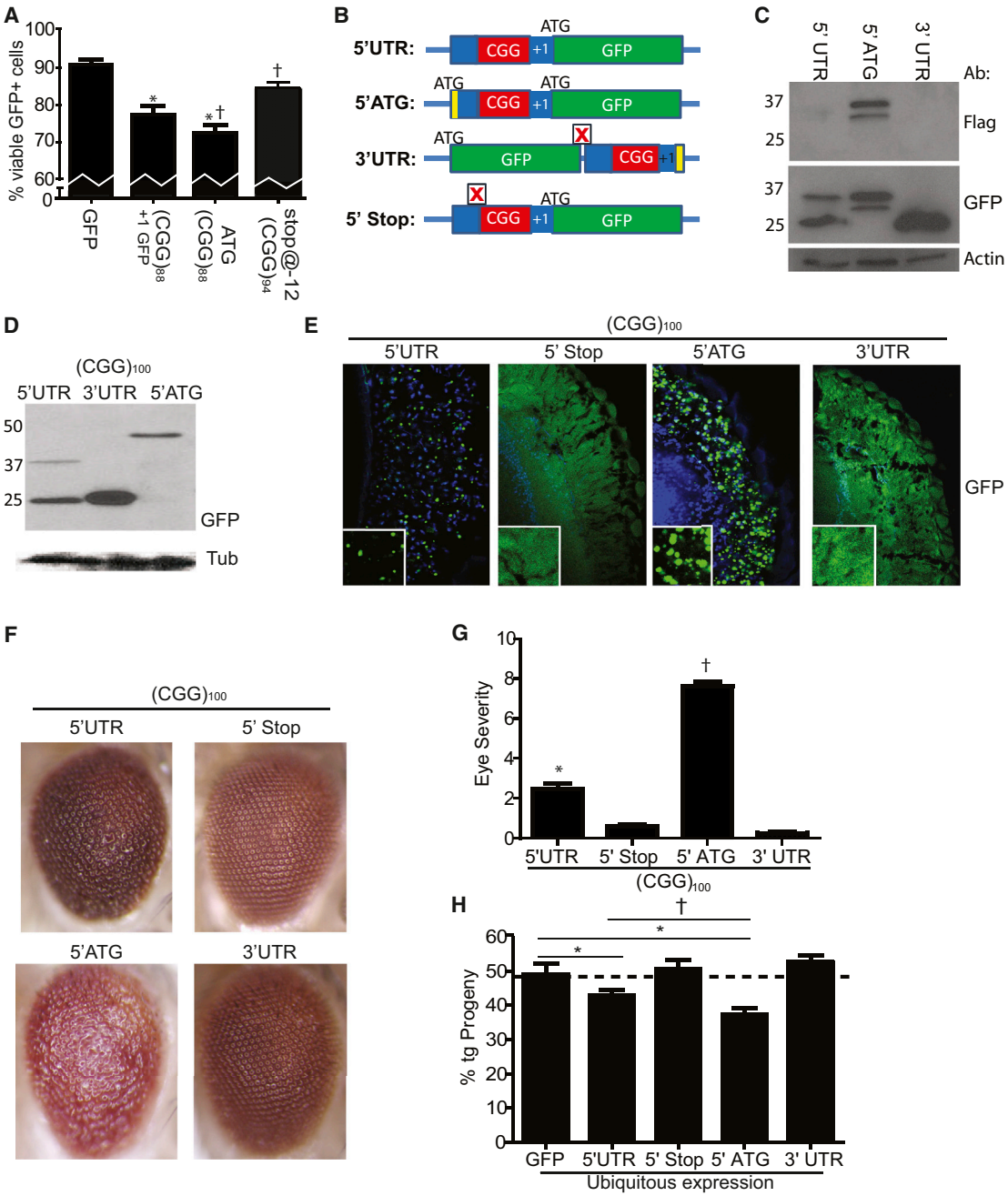


Figure 7. Production of Polyglycine Protein from CGG Repeat RNA Constructs Correlates with Toxicity in *Drosophila* Models

(A) COS cell viability 72 hr after transfection of (from left to right) GFP alone, $(CGG)_{88}$ +1 (Gly) GFP, ATG-FLAG-$(CGG)_{88}$ +1 (Gly) GFP, or Stop$^{@-12}(CGG)_{94}$ +1 (Gly) GFP. *p < 0.05 versus GFP alone; †p < 0.05 versus $(CGG)_{88}$ GFP.

(B) Schematic of pUAST constructs used to generate fly lines with differing amounts of polyglycine protein production but identical (CGG) repeat RNA expansions. Full sequences are shown in Table S2. Boxed red X, stop codon; green, GFP; red, CGG repeat; blue, other sequence; yellow, epitope tag (Flag or 6xHis tag).

(C) Placing the CGG repeat in the 3′ UTR of GFP eliminates the RAN translation product. COS cells were transiently transfected with either GFP (− control), ATG-FLAG-$(CGG)_{100}$ +1 GFP (+ control), or GFP-STOP-$(CGG)_{100}$-FLAG and total lysates were harvested. Blots were serially probed with antibodies to FLAG, GFP, and actin.

(D) Lysates from *Drosophila* lines expressing the described constructs were probed with GFP or Tubulin. The higher MW of the ATG construct results from 150 nt of intervening sequence between the start codon and the start of the $(CGG)_{100}$ repeat.

(E) GFP inclusion formation in *Drosophila* eye cross-sections of the indicated genotypes.

(F) Representative images of eyes 1–2 days after eclosion with the indicated genotypes demonstrate differential toxicity across lines that is dependent on polyglycine translation.

(legend continued on next page)

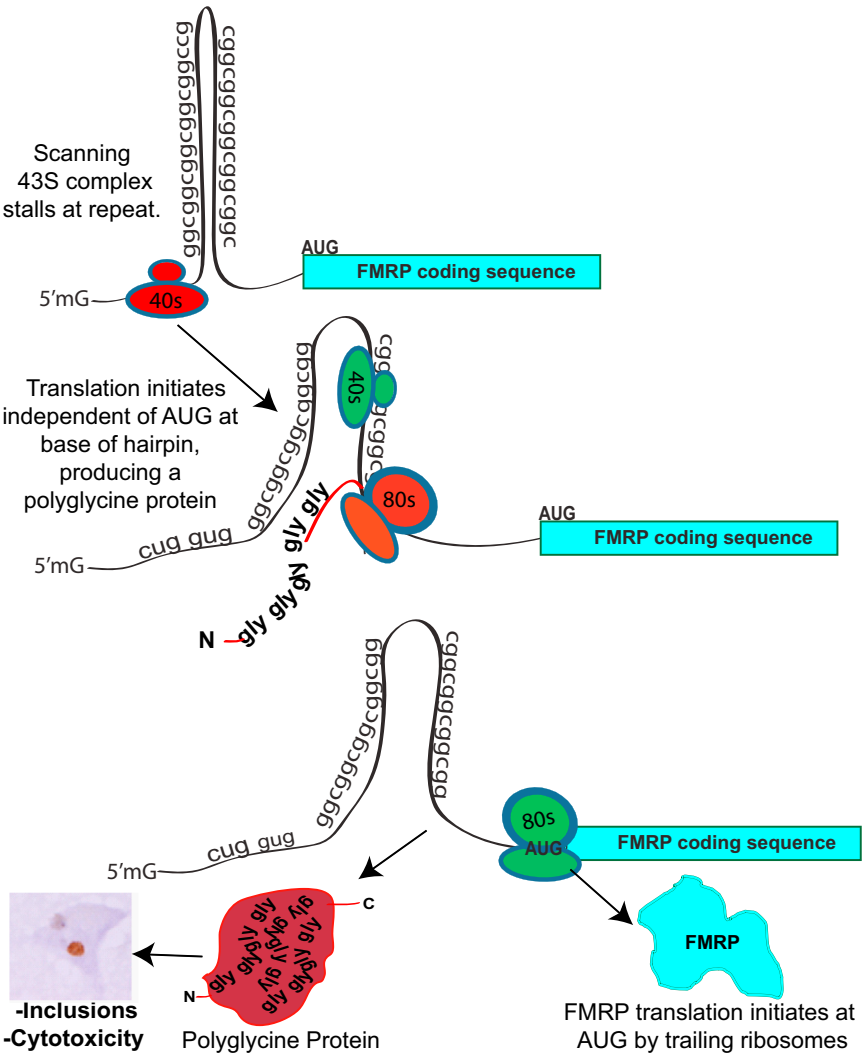


Figure 8. Model for CGG RAN Translation in Fragile X-Associated Tremor Ataxia Syndrome

Ribosomes assemble on the 5′ end of the FMR1 message and scan the mRNA for an appropriate initiation sequence. Near the CGG repeat hairpin, the 43S preinitiaton complex (red) stalls, triggering RAN translational initiation at a near-AUG codon. Once translation initiates, the ribosome reads through the repeat to produce a polyglycine-containing protein. Normally, this peptide is readily cleared from cells, but with larger repeats the resultant expanded polyglycine protein accumulates in inclusions. The downstream AUG start site for FMRP is not in frame with the polyglycine protein; thus, no N-terminal addition onto FMRP occurs with this CGG RAN translation. Trailing ribosomes (green) may not stall at the hairpin but instead initiate translation normally at the AUG for FMRP.

CAG repeats in SCA 8, in which the surrounding sequence appeared to be an important modulator (Zu et al., 2011). Thus, RAN translation may not result from a single mechanism. Rather, each repeat, and indeed each reading frame within each repeat, may have different contextual requirements. These differences notwithstanding, the fact that atypical translation has now been observed independently with four different nucleotide RNA repeats in cell lines, animal models, and human tissues suggests that it is a more widespread biological event than anticipated.

An emerging question now is what roles do these translation initiation events play in normal physiology and in disease? Our findings support a significant role for the FMRpolyG protein in disease pathogenesis, given the evidence in *Drosophila* and mammalian cells of enhanced toxicity with increased polyglycine translation and lessened toxicity when translation is reduced. However, numerous published studies support a primary role for CGG RNA in toxicity (Arocena et al., 2005; Hashem et al., 2009; Jin et al., 2007; Sellier et al., 2010; Sofola et al., 2007), leading us to suggest that, in FXTAS, additive or synergistic toxicity associated with both toxic RNA and toxic proteins may be critical to disease pathogenesis. Though we have focused on FMRpolyG production, other CGG RAN translation-associated products such as the polyalanine protein could represent additional toxic species. Moreover, if RAN translation occurs with CCG repeats, then production of other homopolymeric proteins from the antisense transcript through the repeat could also be relevant (Ladd et al., 2007). For all of these potentially toxic entities, it will be important to determine their relative production in patients and relative degree of toxicity in animal models to ascertain their roles in disease pathogenesis.

CGG RAN translation may also play a normal physiological role in translational regulation of FMR1 mRNA. FMRP, the protein product of *FMR1*, critically regulates synaptic function and its loss leads to fragile X syndrome, a common cause of autism and mental retardation. FMR1 mRNA is rapidly translated at synapses in an activity-dependent manner, where it constrains local synaptic protein translation (Penagarikano et al., 2007). Our results in transfected cells show that RAN translation can

(G) Quantification of the rough-eye phenotypes associated with the indicated genotypes demonstrates significantly enhanced toxicity in lines with enhanced polyglycine translation and reduced toxicity in lines in which CGG RAN translation is blocked.
(H) Ubiquitous expression of CGG GFP transgenes in the 5′ UTR decreases fly viability as measured by progeny eclosion ratios. This change in viability is blocked in CGG 5′ stop lines and CGG 3′ UTR lines but the constructs are more toxic in ATG-CGG GFP lines in which polyglycine protein production is enhanced. For (G) and (H), *p < 0.05 versus 3′ UTR or GFP; †p < 0.05 versus 5′ UTR. Error bars represent SEM.

occur at normal repeat sizes, with initiation occurring within a narrow region just 5′ of the repeat (Figure 8). Analysis of ribosomal profiling data sets derived from samples with normal CGG repeat sizes demonstrates the presence of assembled ribosomes over these regions in both human and mouse cell lines (Figures 5 and S5; Ingolia et al., 2011). Intriguingly, the *Drosophila* homolog of FMRP, *dfxr*, is expressed as two isoforms, with the larger isoform initiating translation at a CUG codon upstream of the canonical TIS, indicating that aspects of this process may be evolutionarily conserved (Beerman and Jongens, 2011). Upstream ORFs are believed to suppress expression from downstream canonical ORFs (Chatterjee and Pal, 2009). In the case of FMR1 mRNA, translation through the repeat may assist RNA unwinding via helicase recruitment, allowing normal scanning by trailing ribosomes and appropriate initiation at the canonical ORF (Figure 8). Alternatively, ribosomes translating through the repeat could terminate translation and reinitiate at the AUG of FMRP, or ribosomes could initiate downstream of the repeat via an internal ribosomal entry site (Ludwig et al., 2011).

Variations on the RAN translation described here potentially could expand the percentage of the transcriptome encoding for protein, complicating the classical definitions by which we divide ''coding'' from ''noncoding'' RNA. Consistent with this, recent unbiased methods in yeast and mammalian cells reveal that thousands of transcripts initiate translation at non-AUG start sites, often creating upstream ORFs in sequences previously identified as 5′ UTR (Ingolia et al., 2011; Ivanov et al., 2011). Usage of these atypical upstream ORFs is responsive to changes in cell state and external stimuli (Brar et al., 2012). Mechanisms similar to those reported here may therefore have broader repercussions for the neuronal proteome and global translational regulation.

EXPERIMENTAL PROCEDURES

Fly Stocks

Drosophila lines used in Figures 1 and S1 of this study have been previously described in Jin et al. (2003), Pandey et al. (2007), and Todd et al. (2010). Details of construction of new fly lines are in the Supplemental Experimental Procedures and Table S2. Unless stated otherwise, all crosses were done and maintained on standard food at 25°C. For $(CGG)_{90}$ GFP lines, stability of the CGG repeat was confirmed by PCR and sequencing using C and F primers.

Plasmid Constructs for Cell Culture Experiments

Cell culture expression plasmids were derived from CMV-$(CGG)_{88}$-GFP, a kind gift from Paul Hagerman (Arocena et al., 2005) or were PCR cloned from patient-derived cell lines. Sequence variants were generated from this vector by site-directed mutagenesis (Stratagene) according to manufacturer's protocols. All vector sequences were confirmed by Sanger sequencing and are described in detail in Table S1.

Lysostaphin Protease Digestion

Immunoprecipitation with anti-GFP agarose beads was conducted as described in the Supplemental Experimental Procedures. Prior to elution, the beads were washed 2× in RIPA without protease inhibitors and then in 20 mM Tris-HCl buffer at pH 7.5 without protease inhibitors. The agarose beads were then incubated with 0.1 mg/ml Lysostaphin (Sigma) in Tris buffer or in Tris buffer alone with agitation at room temperature for 3 or 20 hr. The beads were then washed 1 × 5 min with RIPA buffer and eluted with Laemmli buffer.

Antibody Generation

Monoclonal mouse antibodies 2J7 and 2C13 were developed commercially (Abmart) against a 10 amino acid peptide, LAGLKRRWRS, in the predicted carboxyl terminus of the FMR1polyG. The epitope is predicted to be present in human patient samples, with a near match (8/10 AA) to the original $(CGG)_{90}$ GFP *Drosophila* lines, but not the newly constructed *Drosophila* lines or the transfection vector constructs used in Figures 2, 3, and 4. The epitope is effectively absent (5/10 AA match) from both mouse KI models. Rabbit polyclonal antibody Ab605 was generated commercially (Rockland) against a larger peptide (GLKRRWRSWWWKCGAP) that overlaps more significantly with the predicted sequence in *Drosophila* (16/17) and in both mouse lines (13/17).

Human Tissue

All human tissues were obtained and distributed under oversight by appropriate institution specific review boards. Detailed methods and descriptions of the patient derived samples are included in the Supplemental Experimental Procedures. Briefly, hippocampal, cerebellar, and frontal cortex tissue from two previously described FXTAS patients (Louis et al., 2006) and age- and sex-matched controls (University of Michigan Alzheimer's Disease Brain Bank) were processed using standard techniques. CGG repeat size was determined in both controls and FXTAS patients by DNA isolation followed by PCR using C and F primers.

Statistical Analysis

For graphs of percentages, error bars represent the 95% confidence interval (CI). For all other graphs, error bars represent SEM. For statistical analyses of nonparametric measurements of viability and inclusions, a chi-squared test was performed. For all other analyses, a one-way ANOVA was performed, with post hoc Dunnet's test for multiple comparisons when applicable. Ribosomal profiling data set analysis is detailed in the Supplemental Experimental Procedures.

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures, two tables, and Supplemental Experimental Procedures and can be found with this article online at http://dx.doi.org/10.1016/j.neuron.2013.03.026.

ACKNOWLEDGMENTS

The authors thank Peng Jin for providing $(CGG)_{90}$ GFP fly lines, Karen Usdin for providing $(CGG)_{120}$ KI ''NIH'' mice, and David Nelson and the Fragile X mutant mouse facility for providing $(CGG)_{160}$ KI ''Dutch'' mice (originally generated by Rob Willemsen in Erasmus, the Netherlands), Paul Hagerman for sharing plasmids and patient tissue samples, and Zhe Han and Jinghan Cheng for fly resources and technical assistance. This work was funded by KNS069809A to P.K.T. NAF postdoctoral fellowship to F.H. and RO1 NS038712, RO1 AG034228, and research funds from the Michigan Alzheimer's Disease Center to H.L.P.

Accepted: March 26, 2013
Published: April 18, 2013

REFERENCES

Arocena, D.G., Iwahashi, C.K., Won, N., Beilina, A., Ludwig, A.L., Tassone, F., Schwartz, P.H., and Hagerman, P.J. (2005). Induction of inclusion formation and disruption of lamin A/C structure by premutation CGG-repeat RNA in human cultured neural cells. Hum. Mol. Genet. *14*, 3661–3671.

Ash, P.E., Bieniek, K.F., Gendron, T.F., Caulfield, T., Lin, W.L., Dejesus-Hernandez, M., van Blitterswijk, M.M., Jansen-West, K., Paul, J.W., 3rd, Rademakers, R., et al. (2013). Unconventional translation of C9ORF72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. Neuron *77*, 639–646.

Beerman, R.W., and Jongens, T.A. (2011). A non-canonical start codon in the Drosophila fragile X gene yields two functional isoforms. Neuroscience *181*, 48–66.

Brar, G.A., Yassour, M., Friedman, N., Regev, A., Ingolia, N.T., and Weissman, J.S. (2012). High-resolution view of the yeast meiotic program revealed by ribosome profiling. Science *335*, 552–557.

Brouwer, J.R., Huizer, K., Severijnen, L.A., Hukema, R.K., Berman, R.F., Oostra, B.A., and Willemsen, R. (2008). CGG-repeat length and neuropathological and molecular correlates in a mouse model for fragile X-associated tremor/ataxia syndrome. J. Neurochem. *107*, 1671–1682.

Chatterjee, S., and Pal, J.K. (2009). Role of 5′- and 3′-untranslated regions of mRNAs in human diseases. Biol. Cell *101*, 251–262.

Chen, L.S., Tassone, F., Sahota, P., and Hagerman, P.J. (2003). The (CGG)n repeat element within the 5′ untranslated region of the FMR1 message provides both positive and negative cis effects on in vivo translation of a downstream reporter. Hum. Mol. Genet. *12*, 3067–3074.

Cooper, T.A., Wan, L., and Dreyfuss, G. (2009). RNA and disease. Cell *136*, 777–793.

DeJesus-Hernandez, M., Mackenzie, I.R., Boeve, B.F., Boxer, A.L., Baker, M., Rutherford, N.J., Nicholson, A.M., Finch, N.A., Flynn, H., Adamson, J., et al. (2011). Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. Neuron *72*, 245–256.

Entezam, A., Biacsi, R., Orrison, B., Saha, T., Hoffman, G.E., Grabczyk, E., Nussbaum, R.L., and Usdin, K. (2007). Regional FMRP deficits and large repeat expansions into the full mutation range in a new Fragile X premutation mouse model. Gene *395*, 125–134.

Fritsch, C., Herrmann, A., Nothnagel, M., Szafranski, K., Huse, K., Schumann, F., Schreiber, S., Platzer, M., Krawczak, M., Hampe, J., and Brosch, M. (2012). Genome-wide search for novel human uORFs and N-terminal protein extensions using ribosomal footprinting. Genome Res. *22*, 2208–2218.

Greco, C.M., Berman, R.F., Martin, R.M., Tassone, F., Schwartz, P.H., Chang, A., Trapp, B.D., Iwahashi, C., Brunberg, J., Grigsby, J., et al. (2006). Neuropathology of fragile X-associated tremor/ataxia syndrome (FXTAS). Brain *129*, 243–255.

Guo, H., Ingolia, N.T., Weissman, J.S., and Bartel, D.P. (2010). Mammalian microRNAs predominantly act to decrease target mRNA levels. Nature *466*, 835–840.

Handa, V., Goldwater, D., Stiles, D., Cam, M., Poy, G., Kumari, D., and Usdin, K. (2005). Long CGG-repeat tracts are toxic to human cells: implications for carriers of Fragile X premutation alleles. FEBS Lett. *579*, 2702–2708.

Hashem, V., Galloway, J.N., Mori, M., Willemsen, R., Oostra, B.A., Paylor, R., and Nelson, D.L. (2009). Ectopic expression of CGG containing mRNA is neurotoxic in mammals. Hum. Mol. Genet. *18*, 2443–2451.

Hsieh, A.C., Liu, Y., Edlind, M.P., Ingolia, N.T., Janes, M.R., Sher, A., Shi, E.Y., Stumpf, C.R., Christensen, C., Bonham, M.J., et al. (2012). The translational landscape of mTOR signalling steers cancer initiation and metastasis. Nature *485*, 55–61.

Huber, T.W., and Schuhardt, V.T. (1970). Lysostaphin-induced, osmotically fragile Staphylococcus aureus cells. J. Bacteriol. *103*, 116–119.

Ingolia, N.T., Lareau, L.F., and Weissman, J.S. (2011). Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. Cell *147*, 789–802.

Ingolia, N.T., Brar, G.A., Rouskin, S., McGeachy, A.M., and Weissman, J.S. (2012). The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. Nat. Protoc. *7*, 1534–1550.

Ivanov, I.P., Firth, A.E., Michel, A.M., Atkins, J.F., and Baranov, P.V. (2011). Identification of evolutionarily conserved non-AUG-initiated N-terminal extensions in human coding sequences. Nucleic Acids Res. *39*, 4220–4234.

Iwahashi, C.K., Yasui, D.H., An, H.J., Greco, C.M., Tassone, F., Nannen, K., Babineau, B., Lebrilla, C.B., Hagerman, R.J., and Hagerman, P.J. (2006). Protein composition of the intranuclear inclusions of FXTAS. Brain *129*, 256–271.

Jacquemont, S., Hagerman, R.J., Leehey, M.A., Hall, D.A., Levine, R.A., Brunberg, J.A., Zhang, L., Jardini, T., Gane, L.W., Harris, S.W., et al. (2004). Penetrance of the fragile X-associated tremor/ataxia syndrome in a premutation carrier population. JAMA *291*, 460–469.

Jin, P., Zarnescu, D.C., Zhang, F., Pearson, C.E., Lucchesi, J.C., Moses, K., and Warren, S.T. (2003). RNA-mediated neurodegeneration caused by the fragile X premutation rCGG repeats in Drosophila. Neuron *39*, 739–747.

Jin, P., Duan, R., Qurashi, A., Qin, Y., Tian, D., Rosser, T.C., Liu, H., Feng, Y., and Warren, S.T. (2007). Pur alpha binds to rCGG repeats and modulates repeat-mediated neurodegeneration in a Drosophila model of fragile X tremor/ataxia syndrome. Neuron *55*, 556–564.

Ladd, P.D., Smith, L.E., Rabaia, N.A., Moore, J.M., Georges, S.A., Hansen, R.S., Hagerman, R.J., Tassone, F., Tapscott, S.J., and Filippova, G.N. (2007). An antisense transcript spanning the CGG repeat region of FMR1 is upregulated in premutation carriers but silenced in full mutation individuals. Hum. Mol. Genet. *16*, 3174–3187.

Lee, S., Liu, B., Lee, S., Huang, S.X., Shen, B., and Qian, S.B. (2012). Global mapping of translation initiation sites in mammalian cells at single-nucleotide resolution. Proc. Natl. Acad. Sci. USA *109*, E2424–E2432.

Leehey, M.A., Berry-Kravis, E., Goetz, C.G., Zhang, L., Hall, D.A., Li, L., Rice, C.D., Lara, R., Cogswell, J., Reynolds, A., et al. (2008). FMR1 CGG repeat length predicts motor dysfunction in premutation carriers. Neurology *70*, 1397–1402.

Li, L.B., Yu, Z., Teng, X., and Bonini, N.M. (2008). RNA toxicity is a component of ataxin-3 degeneration in Drosophila. Nature *453*, 1107–1111.

Louis, E., Moskowitz, C., Friez, M., Amaya, M., and Vonsattel, J.P. (2006). Parkinsonism, dysautonomia, and intranuclear inclusions in a fragile X carrier: a clinical-pathological study. Mov. Disord. *21*, 420–425.

Ludwig, A.L., Hershey, J.W., and Hagerman, P.J. (2011). Initiation of translation of the FMR1 mRNA Occurs predominantly through 5′-end-dependent ribosomal scanning. J. Mol. Biol. *407*, 21–34.

Mori, K., Weng, S.M., Arzberger, T., May, S., Rentzsch, K., Kremmer, E., Schmid, B., Kretzschmar, H.A., Cruts, M., Van Broeckhoven, C., et al. (2013). The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTLD/ALS. Science *339*, 1335–1338.

Moseley, M.L., Zu, T., Ikeda, Y., Gao, W., Mosemiller, A.K., Daughters, R.S., Chen, G., Weatherspoon, M.R., Clark, H.B., Ebner, T.J., et al. (2006). Bidirectional expression of CUG and CAG expansion transcripts and intranuclear polyglutamine inclusions in spinocerebellar ataxia type 8. Nat. Genet. *38*, 758–769.

Napierala, M., Michalowski, D., de Mezer, M., and Krzyzosiak, W.J. (2005). Facile FMR1 mRNA structure regulation by interruptions in CGG repeats. Nucleic Acids Res. *33*, 451–463.

Oma, Y., Kino, Y., Sasagawa, N., and Ishiura, S. (2004). Intracellular localization of homopolymeric amino acid-containing proteins expressed in mammalian cells. J. Biol. Chem. *279*, 21217–21222.

Orr, H.T., and Zoghbi, H.Y. (2007). Trinucleotide repeat disorders. Annu. Rev. Neurosci. *30*, 575–621.

Pandey, U.B., Nie, Z., Batlevi, Y., McCray, B.A., Ritson, G.P., Nedelsky, N.B., Schwartz, S.L., DiProspero, N.A., Knight, M.A., Schuldiner, O., et al. (2007). HDAC6 rescues neurodegeneration and provides an essential link between autophagy and the UPS. Nature *447*, 859–863.

Pearson, C.E. (2011). Repeat associated non-ATG translation initiation: one DNA, two transcripts, seven reading frames, potentially nine toxic entities! PLoS Genet. *7*, e1002018.

Penagarikano, O., Mulle, J.G., and Warren, S.T. (2007). The pathophysiology of fragile x syndrome. Annu. Rev. Genomics Hum. Genet. *8*, 109–129.

Renton, A.E., Majounie, E., Waite, A., Simón-Sánchez, J., Rollinson, S., Gibbs, J.R., Schymick, J.C., Laaksovirta, H., van Swieten, J.C., Myllykangas, L., et al.; ITALSGEN Consortium. (2011). A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. Neuron *72*, 257–268.

Sellier, C., Rau, F., Liu, Y., Tassone, F., Hukema, R.K., Gattoni, R., Schneider, A., Richard, S., Willemsen, R., Elliott, D.J., et al. (2010). Sam68 sequestration and partial loss of function are associated with splicing alterations in FXTAS patients. EMBO J. *29*, 1248–1261.

Sofola, O.A., Jin, P., Qin, Y., Duan, R., Liu, H., de Haro, M., Nelson, D.L., and Botas, J. (2007). RNA-binding proteins hnRNP A2/B1 and CUGBP1 suppress fragile X CGG premutation repeat-induced neurodegeneration in a Drosophila model of FXTAS. Neuron *55*, 565–571.

Stochmanski, S.J., Therrien, M., Laganière, J., Rochefort, D., Laurent, S., Karemera, L., Gaudet, R., Vyboh, K., Van Meyel, D.J., Di Cristo, G., et al. (2012). Expanded ATXN3 frameshifting events are toxic in Drosophila and mammalian neuron models. Hum. Mol. Genet. *21*, 2211–2218.

Tassone, F., Iwahashi, C., and Hagerman, P.J. (2004). FMR1 RNA within the intranuclear inclusions of fragile X-associated tremor/ataxia syndrome (FXTAS). RNA Biol. *1*, 103–105.

Thoreen, C.C., Chantranupong, L., Keys, H.R., Wang, T., Gray, N.S., and Sabatini, D.M. (2012). A unifying model for mTORC1-mediated regulation of mRNA translation. Nature *485*, 109–113.

Todd, P.K., Oh, S.Y., Krans, A., Pandey, U.B., Di Prospero, N.A., Min, K.T., Taylor, J.P., and Paulson, H.L. (2010). Histone deacetylases suppress CGG repeat-induced neurodegeneration via transcriptional silencing in models of fragile X tremor ataxia syndrome. PLoS Genet. *6*, e1001240.

Wilburn, B., Rudnicki, D.D., Zhao, J., Weitz, T.M., Cheng, Y., Gu, X., Greiner, E., Park, C.S., Wang, N., Sopher, B.L., et al. (2011). An antisense CAG repeat transcript at JPH3 locus mediates expanded polyglutamine protein toxicity in Huntington's disease-like 2 mice. Neuron *70*, 427–440.

Willemsen, R., Hoogeveen-Westerveld, M., Reis, S., Holstege, J., Severijnen, L.A., Nieuwenhuizen, I.M., Schrier, M., van Unen, L., Tassone, F., Hoogeveen, A.T., et al. (2003). The FMR1 CGG repeat mouse displays ubiquitin-positive intranuclear neuronal inclusions; implications for the cerebellar tremor/ataxia syndrome. Hum. Mol. Genet. *12*, 949–959.

Williams, A.J., and Paulson, H.L. (2008). Polyglutamine neurodegeneration: protein misfolding revisited. Trends Neurosci. *31*, 521–528.

Zu, T., Gibbens, B., Doty, N.S., Gomes-Pereira, M., Huguet, A., Stone, M.D., Margolis, J., Peterson, M., Markowski, T.W., Ingram, M.A., et al. (2011). Non-ATG-initiated translation directed by microsatellite expansions. Proc. Natl. Acad. Sci. USA *108*, 260–265.

ACTA NEUROPATHOLOGICA COMMUNICATIONS

LETTER TO THE EDITOR Open Access

FMRpolyG-positive inclusions in CNS and non-CNS organs of a fragile X premutation carrier with fragile X-associated tremor/ataxia syndrome

Ronald AM Buijsen[1], Chantal Sellier[2], Lies-Anne WFM Severijnen[1], Mustapha Oulad-Abdelghani[2], Rob FM Verhagen[1], Robert F Berman[3], Nicolas Charlet-Berguerand[2], Rob Willemsen[1†] and Renate K Hukema[1*†]

Keywords: FXTAS, CGG repeat, FMRpolyG, RAN translation, Gain-of-function, Inclusions

Fragile X-associated Tremor/Ataxia syndrome (FXTAS), a late-onset monogenetic neurodegenerative disorder, is caused by a CGG-repeat expansion (55-200) in the 5′ UTR of the fragile-X mental retardation 1 gene (*FMR1*) on the X-chromosome [1]. The prevalence of the *FMR1* premutation (PM) is about 1:855 in males and 1:291 in females [2]. Approximately 45.5% of male and 16.5% of female PM carriers older than 50 years will develop signs of FXTAS [3]. In addition to the core features of tremor and gait ataxia, unexplained medical co-morbidities have been reported, including thyroid disease, cardiac arrhythmias, hypertension, migraine, impotence, and neuropathy [4]. PM carriers have increased levels of *FMR1* mRNA (2 to 8 fold in leucocytes) and normal to slightly reduced FMR1 protein (FMRP) levels [5]. The current hypothesis is that FXTAS is caused by an RNA gain-of-function mechanism. Ubiquitin-positive intranuclear inclusions, are found in both brain and non-central nervous system (CNS) organs of patients with FXTAS [6,7]. So far, it is not clear whether these inclusions are protective or toxic. Recently, it has been hypothesized that repeat-associated non-AUG (RAN) translation plays a role in disease process and inclusion formation. Todd et al. [8] demonstrated that through initiation at a near-ATG codon located in the 5′UTR of the *FMR1* gene a polyGlycine-containing protein, FMRpolyG, is expressed. This protein accumulates in ubiquitin-positive inclusions in *Drosophila*, cell culture, mouse disease models and brain from FXTAS patients. To investigate the link between FMRpolyG expression and the co-morbid medical problems associated with the PM we have developed two novel mouse monoclonal antibodies against polyGlycine; 8FM and 9FM (for epitopes and specificity see Additional file 1: Figure S1), and performed immunostaining in CNS as well as in non-CNS organs of FXTAS patient J.L. (case 6 in [7]; other cases not available). To establish antibody specificity, we performed immunostaining with both antibodies on brain sections from FXTAS patient J.L., healthy non-demented controls (n = 3) and a patient with Parkinson disease, Alzheimer disease, or C9FTD. In hippocampus and cerebellum from FXTAS patient J.L. we identified FMRpolyG-positive inclusions with both 8FM (1:10) and 9FM (1:10) antibody (Figure 1a-b, Additional file 2: Figure S2a-b), as was described previously [8]. None of the controls showed FMRpolyG-positive inclusions (data not shown). Next, we studied the immunolocalization of FMRpolyG protein in heart, kidney, adrenal gland and thyroid in patient J.L. with 8FM (1:10) and 9FM (1:10), compared to post mortem non-CNS somatic organ tissues from 3 healthy controls. We also examined tissues for FMRP (mouse T1A; 1:200) expression and ubiquitin-positive inclusions (DAKO, ZO458; 1:200). Consistent with our previous report [7], ubiquitin-positive intranuclear inclusions were identified along with a normal distribution of FMRP (data not shown). Intranuclear FMRpolyG-positive inclusions could be detected in all organs examined (Figure 1c-h, Additional file 2: Figure S2c-h). No control tissues showed any FMRpolyG-positive inclusions (data not shown). Colocalization of ubiquitin- and FMRpolyG-positive inclusions was visualized and quantified by immunofluorescent double staining using antibodies against ubiquitin and FMRpolyG

* Correspondence: r.hukema@erasmusmc.nl
†Equal contributors
[1]Department of Clinical Genetics, Erasmus, MC, PO Box 2040, 3000CA, The Netherlands
Full list of author information is available at the end of the article

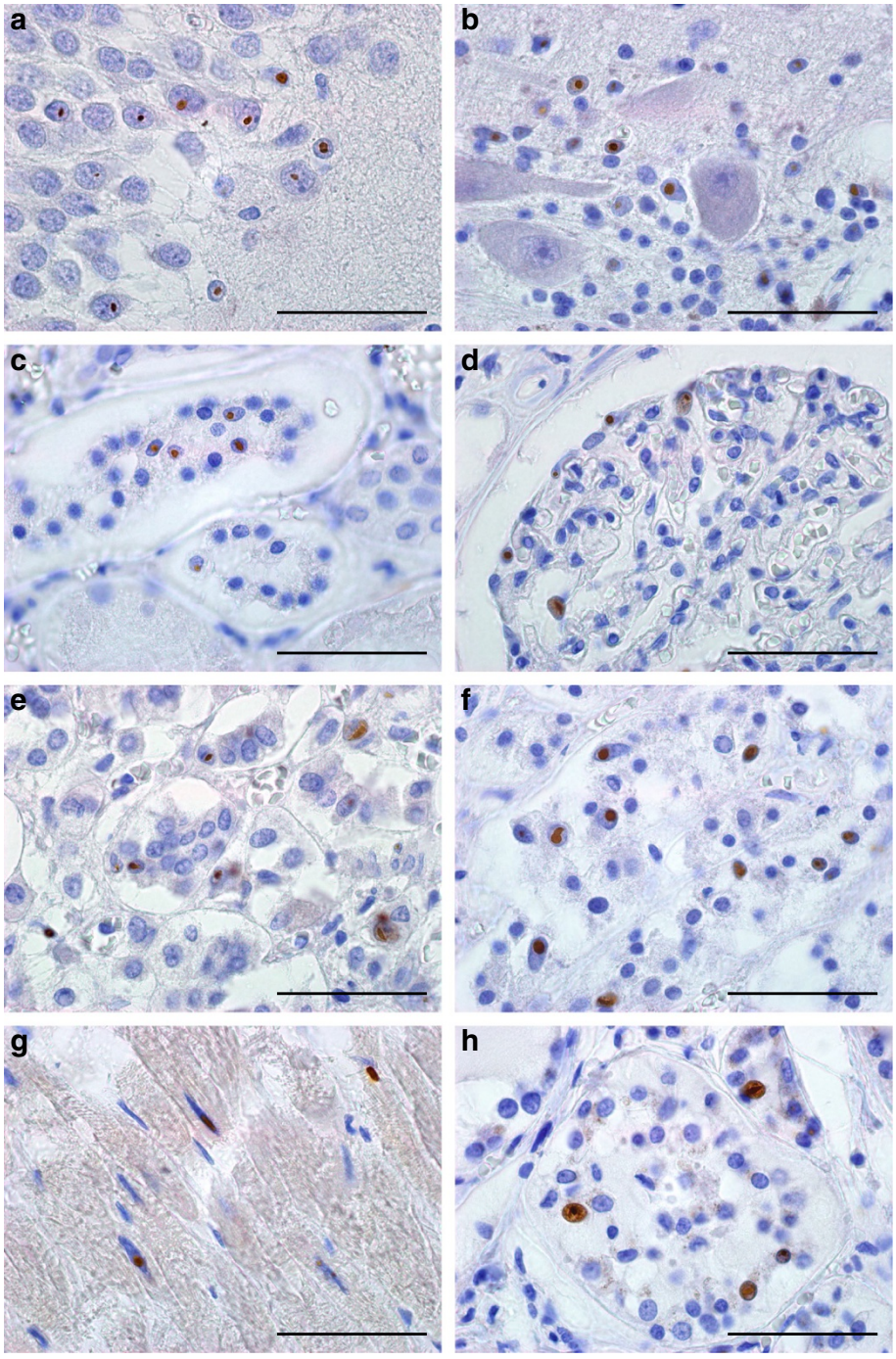


Figure 1 9FM FMRpolyG-positive intranuclear inclusions in hippocampus, cerebellum and non-CNS tissues of a FXTAS patient. FMRpolyG-positive (9FM) intranuclear inclusions in **a** hippocampus**, b** cerebellum, **c** glomeruli and **d** distal tubule of the kidney, **e** zona glomerulosa and **f** zona reticularis of adrenal gland, **g** cardiomyocytes and **h** thyroid. All sections were immunostained with 9FM antibody and counterstained with hematoxylin. Scale bars represent 50 μm.

(8FM) (Figure 2a-f). For hippocampus, cerebellum and the non-CNS organs most inclusions are positive for both FMRpolyG and ubiquitin, although some rare inclusions positive for only one of the proteins could also be detected (Figure 2g, n = 100 inclusions). In conclusion, using two novel antibodies the present report not only confirms the existence of FMRpolyG-positive aggregates in CNS tissue from a FXTAS individual but also demonstrates for the first time the presence of FMRpolyG-positive intranuclear inclusions in post mortem non-CNS material of a PM carrier

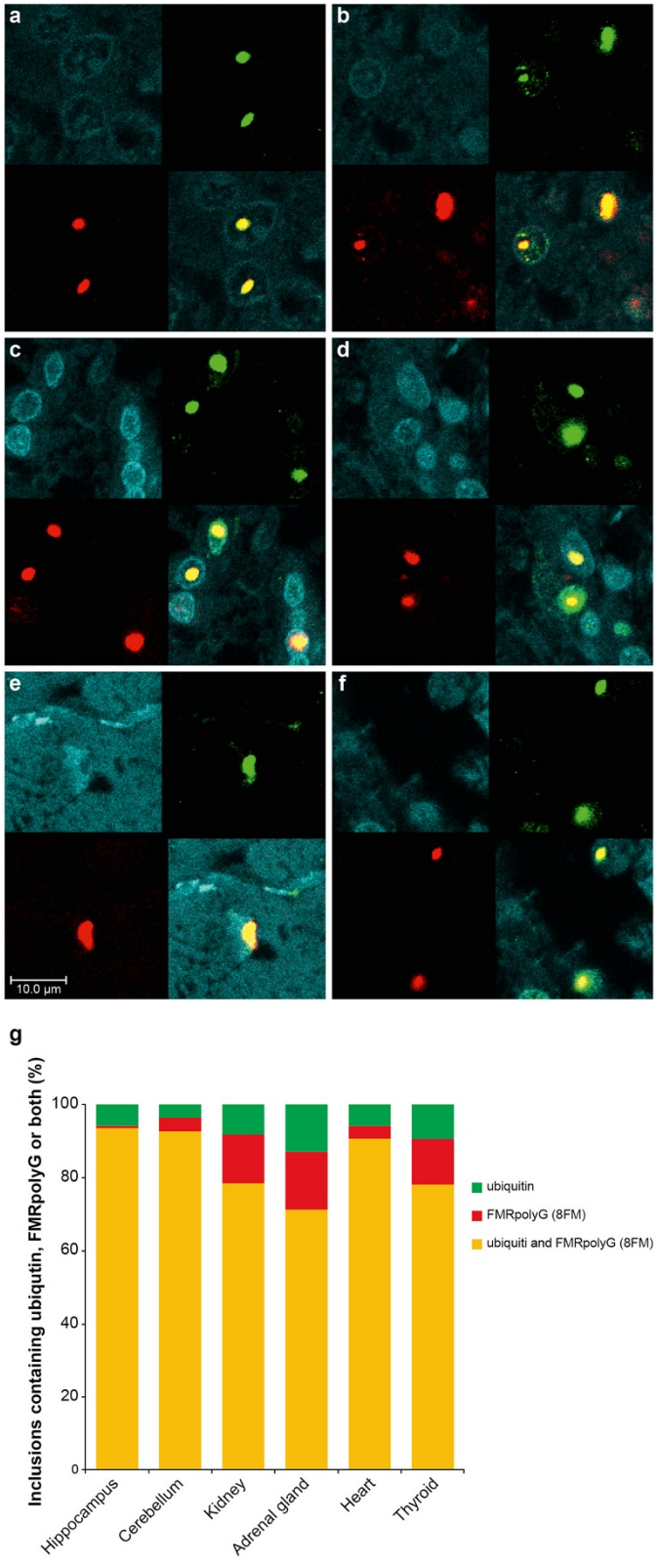


Figure 2 (See legend on next page.)

(See figure on previous page.)
Figure 2 Colocalization of FMRpolyG (8FM) and ubiquitin in intranuclear inclusions in hippocampus, cerebellum and of non-CNS tissues of a FXTAS patient. Staining for ubiquitin (green), FMRpolyG (8FM; red) and DAPI (blue). Colocalization of ubiquitin and FMRpolyG (yellow) is seen in **a** hippocampus**, b** cerebellum, **c** kidney, **d** adrenal gland, **e** cardiomyocytes, and **f** thyroid; **g** quantification of inclusions containing ubiquitin and/or FMRpolyG (n = 100). Scale bars represent 10 μm.

with FXTAS. Furthermore, colocalization of FMRpolyG and ubiquitin is found in the vast majority of inclusions. The presence of FMRpolyG-positive intranuclear inclusions in heart, kidney, adrenal gland and thyroid is consistent with the unexplained medical co-morbidities reported in some patients with FXTAS, including thyroid disease, cardiac arrhythmias, hypertension, migraine, impotence, and neuropathy. We hypothesize that the underlying pathological mechanisms of the medical co-morbidities in systemic tissues share common features (protein toxic gain-of-function) with CNS pathology of patients with FXTAS. Our report suggests that in addition to elevated levels of *FMR1* mRNA containing an expanded CGG repeat, and ubiquitin-positive inclusions, FMRpolyG expression might also play a role in a toxic gain-of-function mechanism in medical co-morbidities in FXTAS (RNA versus FMRpolyG toxic gain-of-function). Interestingly, a very recent report suggests that RAN translation products in C9FTD/ALS, toxic dipeptide repeat proteins (poly-(glycine-arginine) and poly-(proline-arginine)), are toxic in *Drosophila* [9]. Further research is needed to understand how FMRpolyG may elicit toxicity in both CNS and non-CNS organs and its precise role in co-morbidities in PM carriers. Importantly, if FMRpolyG production is important for cellular toxicity this will open new avenues for therapeutic intervention studies for FXTAS by developing drugs that block this aberrant translation.

Additional files

Additional file 1: Figure S1. Epitopes and specificity FMRpolyG antibodies 8FM and 9FM. a Sequence of the 5′UTR of human *FMR1* gene and the FMRpolyG peptide sequence resulting from RAN translation. Epitopes of 8FM and 9FM antibodies are boxed. b Monoclonal 8FM and 9FM antibodies were validated on COS7 cells transfected with pEGFP (Clonetech) plasmid containing the 5′UTR of *FMR1* with 5O CGG repeats, thus expressing the polyGlycine protein fused to GFP. On Western Blot a specific product could be detected for both antibodies and no product was detectable in COS7 cells transfected with a control GFP plasmid. c As an additional control experiment we performed immunostainings for 8FM antibody and could demonstrate specific intranuclear inclusions in COS7 cells transfected with a construct expressing the FMRpolyG fused to GFP, while cells tranfected with only GFP did not show any inclusion formation. Identical results were obtained with 9FM antibody (data not shown).

Additional file 2: Figure S2. 8FM FMRpolyG-positive intranuclear inclusions in hippocampus, cerebellum and non-CNS tissues of a FXTAS patient. FMRpolyG-positive (8FM) intranuclear inclusions in a hippocampus, b cerebellum, c glomeruli and d distal tubule of the kidney, e zona glomerulosa and f zona reticularis of adrenal gland, g cardiomyocytes and h thyroid. All sections were immunostained with 8FM antibody and counterstained with hematoxylin. Scale bars represent 50 μm.

Abbreviations
ALS: Amyotrophic lateral sclerosis; C9FTD: Chromosome 9 open reading frame 72 frontotemporal dementia; CNS: Central nervous system; *FMR1*: Fragile-X mental retardation 1; FMRP: Fragile-X mental retardation 1 protein; FXTAS: Fragile X-associated Tremor/Ataxia syndrome; PM: Fragile-X mental retardation 1 premutation; RAN: Repeat-associated non-AUG.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
RAMB performed immunostainings and imaging, interpreted results, and drafted the manuscript. CS developed and validated antibodies. LAS performed immunostainings and imaging, and interpreted results. MOA developed and validated antibodies. RFMV performed immunostainings and imaging, and interpreted results. RFB interpreted results, drafted the manuscript and, obtained funding. NCB developed and validated antibodies, interpreted results, drafted the manuscript, and obtained funding. RW interpreted the results, drafted the manuscript, supervised the study, and obtained funding. RKH interpreted the results, drafted the manuscript, supervised the study, and obtained funding. RW and RKH contributed equally to this work. All authors read and approved the final manuscript.

Acknowledgements
The authors wish to acknowledge the contribution of Tom de Vries Lentsch and Nathan Wubben. This work was supported by the Dutch Brain Foundation project number 2012(1)101 and French Muscular Dystrophy Association project number 16649 [to RW], by the National Institutes of Health grant number NINDS NS079775 [to RFB and RW], by E-Rare project number 40-42900-98-1001/113301201 from ZonMW [to RKH], by E-Rare "CURE FXTAS" from ANR [NCB] and ERC "RNA DISEASES" [NCB].

Author details
[1]Department of Clinical Genetics, Erasmus, MC, PO Box 2040, 3000CA, The Netherlands. [2]Department of Neurobiology and Genetics, IGBMC, INSERM U964, CNRS UMR7104, University of Strasbourg, Illkirch, France. [3]Department of Neurological Surgery, UC Davis, Davis, CA 95618, USA.

Received: 8 October 2014 Accepted: 8 November 2014

References
1. Hagerman RJ, Leehey M, Heinrichs W, Tassone F, Wilson R, Hills J, Grigsby J, Gage B, Hagerman PJ (2001) Intention tremor, parkinsonism, and generalized brain atrophy in male carriers of fragile X. Neurology 57(1):127–130
2. Hunter J, Rivero-Arias O, Angelov A, Kim E, Fotheringham I, Leal J (2014) Epidemiology of fragile X syndrome: a systematic review and meta-analysis. Am J Med Genet A 164(7):1648–1658
3. Rodriguez-Revenga L, Madrigal I, Pagonabarraga J, Xuncla M, Badenas C, Kulisevsky J, Gomez B, Mila M (2009) Penetrance of FMR1 premutation associated pathologies in fragile X syndrome families. Eur J Hum Genet 17(10):1359–1362
4. Willemsen R, Levenga J, Oostra B (2011) CGG repeat in the FMR1 gene: size matters. Clin Genet 80(3):214–225
5. Tassone F, Hagerman RJ, Taylor AK, Gane LW, Godfrey TE, Hagerman PJ (2000) Elevated levels of FMR1 mRNA in carrier males: a new mechanism of involvement in the Fragile-X syndrome. Am J Hum Genet 66(1):6–15
6. Greco CM, Hagerman RJ, Tassone F, Chudley AE, Del Bigio MR, Jacquemont S, Leehey M, Hagerman PJ (2002) Neuronal intranuclear inclusions in a new cerebellar tremor/ataxia syndrome among fragile X carriers. Brain 125(Pt 8):1760–1771
7. Hunsaker MR, Greco CM, Spath MA, Smits AP, Navarro CS, Tassone F, Kros JM, Severijnen LA, Berry-Kravis EM, Berman RF, Hagerman PJ, Willemsen R,

Hagerman RJ, Hukema RK (2011) Widespread non-central nervous system organ pathology in fragile X premutation carriers with fragile X-associated tremor/ataxia syndrome and CGG knock-in mice. Acta Neuropathol 122:467–479

8. Todd PK, Oh SY, Krans A, He F, Sellier C, Frazer M, Renoux AJ, Chen KC, Scaglione KM, Basrur V, Elenitoba-Johnson K, Vonsattel JP, Louis ED, Sutton MA, Taylor JP, Mills RE, Charlet-Berguerand N, Paulson HL (2013) CGG repeat-associated translation mediates neurodegeneration in fragile X tremor ataxia syndrome. Neuron 78(3):440–455
9. Mizielinska S, Gronke S, Niccoli T, Ridler CE, Clayton EL, Devoy A, Moens T, Norona FE, Woollacott IO, Pietrzyk J, Cleverley K, Nicoll AJ, Pickering-Brown S, Dols J, Cabecinha M, Hendrich O, Fratta P, Fisher EM, Partridge L, Isaacs AM (2014) C9orf72 repeat expansions cause neurodegeneration in Drosophila through arginine-rich proteins. Science 345((6201):1192–1194, doi:10.1126/science.1256800

doi:10.1186/s40478-014-0162-2
Cite this article as: Buijsen *et al.*: FMRpolyG-positive inclusions in CNS and non-CNS organs of a fragile X premutation carrier with fragile X-associated tremor/ataxia syndrome. *Acta Neuropathologica Communications* 2014 **2**:162.

Translation of expanded CGG repeats into FMRpolyG is pathogenic in Fragile X Tremor Ataxia Syndrome.

Chantal Sellier[1#], Ronald AM. Buijsen[2], Fang He[3,4,5], Sam Natla[3,4], Laura Jung[1], Philippe Tropel[1], Angeline Gaucherot[1], Hugues Jacobs[1,6], Hamid Meziane[1,6], Alexandre Vincent[1], Marie-France Champy[1,6], Tania Sorg[1,6], Guillaume Pavlovic[1,6], Marie Wattenhofer-Donze[1,6], Marie-Christine Birling[1,6], Mustapha Oulad-Abdelghani[1], Pascal Eberling[1], Frank Ruffenach[1], Mathilde Joint[1], Mathieu Anheim[7], Veronica Martinez-Cerdeno[8,9], Flora Tassone[9], Rob Willemsen[2], Renate K. Hukema[2], Stéphane Viville[1,10,11], Cecile Martinat[12], Peter K. Todd[3,4], Nicolas Charlet-Berguerand[1#]

1. Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), INSERM U964, CNRS UMR7104, University of Strasbourg, 67400 Illkirch, France.
2. Department of Clinical Genetics, Erasmus MC, 3015 Rotterdam, The Netherlands.
3. Department of Neurology, University of Michigan, Ann Arbor, 48109 Michigan, USA.
4. Veteran Association Health System, Ann Arbor, 48105 Michigan, USA.
5. Department of Biological and Health Sciences, Texas A&M University-Kingsville, USA.
6. PHENOMIN, Institut Clinique de la Souris (ICS), INSERM U964, CNRS UMR7104, University of Strasbourg, 67400 Illkirch, France
7. Department of Neurology, University Hospital of Strasbourg, Hôpital de Hautepierre, Strasbourg, France.
8. Department of Pathology and Laboratory Medicine at UC Davis, Institute for Pediatric Regenerative Medicine, and Shriners Hospitals for Children Northern California.
9. M.I.N.D. Institute, University of California, Davis, Health System, Sacramento, 95817 California, USA.
10. Laboratoire de diagnostic génétique, UF3472 – Infertilité, Nouvel Hôpital Civil, 1 place de l'Hôpital, 67091 Strasbourg, France.
11. IPPTS, 3 rue Koeberlé, 67000 Strasbourg, France.
12. INSERM/UEVE UMR 861, I-STEM, AFM, 91030 Evry, France.

Correspondence to either Chantal Sellier or Nicolas Charlet-Berguerand (lead contact), Department of Translational Medicine and neurogenetics, IGBMC, 1 rue Laurent Fries, 67400 Illkirch, France. Phone: +33 388 653 309, Fax: +33 388 653 201, Email: sellier@igbmc.fr or ncharlet@igbmc.fr

ABSTRACT

Fragile X-associated Tremor/Ataxia Syndrome (FXTAS) is a neurodegenerative disorder caused by a limited expansion of CGG repeats in the 5'UTR of *FMR1*. Two mechanisms are proposed to cause FXTAS: RNA gain-of-function where CGG RNA sequesters specific proteins, and translation of CGG repeats into a polyglycine-containing protein, FMRpolyG. Here, we developed transgenic mice expressing CGG repeats RNA with or without FMRpolyG. Importantly, expression of FMRpolyG is pathogenic, while the sole expression of CGG RNA is not. Toxicity of FMRpolyG is partly mediated by its carboxy-terminus in neuronal cultures and in *Drosophila* models. FMRpolyG interacts with the nuclear lamina protein, LAP2ß and disorganizes the nuclear lamina architecture in neurons differentiated from FXTAS iPS cells. Finally, expression of LAP2ß rescues neuronal death induced by FMRpolyG. Overall, these results suggest that translation of expanded CGG repeats into FMRpolyG, which alters nuclear lamina architecture, drives pathogenesis in FXTAS.

INTRODUCTION

Fragile X-associated Tremor/Ataxia Syndrome (FXTAS) is a neurodegenerative disorder caused by an expansion, called premutation, of 55 to 200 CGG repeats in the 5'untranslated region of the *Fragile X Mental Retardation 1* (*FMR1*) gene located on the X chromosome (Hagerman et al., 2001). The carrier prevalence of the CGG premutation is approximately 1 of ~200 females and ~450 males, but due to incomplete penetrance, it is estimated that 1 in 3,000 men older than 50 years will develop FXTAS (Jacquemont et al., 2004; Seltzer et al., 2012; Tassone et al., 2012). The clinical features of FXTAS include progressive intention tremor and gait ataxia, frequently accompanied by progressive cognitive decline, parkinsonism, peripheral neuropathy and autonomic dysfunctions (Jacquemont et al., 2003). Principal neuropathologies of FXTAS include mild brain atrophy and white matter lesions with the presence of ubiquitin-positive nuclear neuronal and astrocytic inclusions (Greco et al., 2002; Greco et al., 2006). In contrast to Fragile X syndrome, where expanded full mutation alleles (>200 CGG repeats) result in hypermethylation and silencing of the *FMR1* gene, FXTAS carriers of premutation expanded alleles (55–200 CGG repeats) present increased levels of *FMR1* mRNA but slightly reduced expression of the protein encoded by *FMR1*, FMRP (Tassone et al., 2000; Kenneson et al., 2001; Tassone et al., 2007).

Because FXTAS is not observed in Fragile X patients with completely hypermethylated and consequentely fully silenced *FMR1* alleles, a pathogenic mechanism based on expression of mutant *FMR1* mRNAs containing expanded CGG repeats have been proposed (Hagerman et al., 2004). In support of this hypothesis, multiple studies have demonstrated adverse consequences of expressing RNA containing expanded CGG repeats in cell, fly and mouse models (Willemsen et al., 2003; Jin et al., 2003; Arocena et al., 2005; Hashem et al., 2009; Entezam et al., 2007; Hukema et al., 2014; Hukema et al., 2015). However, how *FMR1* mRNA containing expanded CGG repeats is pathogenic is unclear. A first proposed model is that FXTAS results from a toxic RNA gain-of-function mechanism, in which mutant RNA containing expanded CGG repeats would be pathogenic by sequestrating specific RNA binding proteins, ultimately resulting in neuronal cell dysfunctions (Iwahashi et al., 2006; Sofola et al., 2007; Jin et al., 2007; Sellier et al., 2010; Sellier et al., 2013). A second and more recent proposed mechanism is that FXTAS is caused by Repeat-Associated Non-AUG (RAN) translation of the expanded CGG repeats into polyalanine and polyglycine containing proteins, named FMRpolyA and FMRpolyG (Todd et al., 2013). Laura Ranum and colleagues

originally demonstrated that expanded CAG repeats can be translated in all three frames in the absence of any AUG initiation codon (Zu et al., 2011). Subsequently, RAN translation was described and proposed as a causative mechanism in various inherited microsatellite neurodegenerative disorders (reviewed in Cleary and Ranum 2014), including FXTAS (Todd et al., 2013). However, the mechanism and pathological consequences of translating expanded CGG repeats are not fully understood. Specifically, which pathological mechanism (RNA gain-of-function or RAN translation) drives FXTAS pathogenesis is a crucial question.

Here we find that translation of expanded CGG repeats occurs predominantly in the glycine frame through initiation at a near-cognate ACG codon located upstream of the expanded CGG repeats, resulting in expression of FMRpolyG. Next, to define which mechanism drives pathogenicity in FXTAS we developed two novel mouse models, one that expresses both CGG RNA and FMRpolyG and a second one expressing CGG RNA in isolation. Importantly, transgenic mice expressing both CGG RNA repeats and the polyglycine protein (FMRpolyG mouse), but not mice expressing only the mutant RNA containing expanded CGG repeats (CGG RNA mouse) exhibit inclusion formation, motor phenotypes, and reduced lifespan. Deletion analysis of FMRpolyG revealed that its C-terminal region contributes to cell death in mammalian neuronal cultures and decreased viability in *Drosophila*, while the polyglycine stretch promotes protein aggregation. The C-terminus of FMRpolyG interacts with LAP2ß, a protein essential to anchoring lamin proteins to the inner nuclear membrane. Expression of FMRpolyG alters LAP2ß localization, resulting in disorganization of the nuclear lamina architecture, a finding that is also observed in iPS-cell derived neurons from FXTAS patients. Importantly, over-expression of LAP2ß rescues neuronal cell death induced by expression of FMRpolyG. Overall, these results suggest that translation of expanded CGG repeats into FMRpolyG, which disrupts nuclear lamina, is a key pathogenic event in FXTAS.

RESULTS

Translation of expanded CGG repeats initiates at an upstream near-cognate codon.

To confirm a previous observation of translation of expanded CGG repeats in absence of a canonical AUG start codon in *Drosophila* (Todd et al., 2013), we cloned one hundred CGG repeats embedded within the natural human 5'UTR of *FMR1* fused to the GFP deleted of its ATG and in all three possible frames. These frames were named according to the polypeptide potentially encoded by the expanded CGG repeats, namely glycine, alanine and arginine. Cell transfection and immunoblotting against GFP confirm previous data (Todd et al., 2013), and demonstrates that the 5'UTR of *FMR1* with expanded CGG repeats allows translation of a GFP protein with a ~12 kDa N-terminal extension corresponding to the expanded CGG repeats translated into the glycine frame (Figure 1A and supplemental figure 1A). In contrast, we observed little translation in the alanine or arginine frames. Interestingly, we also observed no translation when the *FMR1* 5'UTR sequence located upstream to the repeats was deleted and the expanded CGG repeats were directly fused to the GFP (Figure 1A). Identical results were observed by imaging GFP fluorescence (supplemental figure 1B). As a further control, treatment with Lysostaphin, a polyglycine endopeptidase, confirmed that expanded CGG repeats are translated into a polyglycine-containing protein (supplemental figure 1C). To confirm the importance of the upstream *FMR1* 5'UTR sequence for CGG repeats translation, we performed various deletions and tested expression of FMRpolyG-GFP by immunoblotting. Consistent with previous reports (Todd et al., 2013; Kearse et al., 2016), deletions of a sequence located upstream of the CGG repeats reduced expression of FMRpolyG (Figure 1B), suggesting that translation of FMRpolyG may start upstream of the CGG repeats.

To further characterize the initiation site of expanded CGG repeats translation, we immunoprecipitated FMRpolyG and determined its N-terminal sequence by proteomic analysis after trypsine digestion. LC-MS/MS spectra revealed initiation to an ACG near-cognate codon located 32 nucleotides upstream of the CGG expansion (Figure 1C). To exclude any bias of digestion, we repeated this experiment but with a construct containing a mutation so that a lysine (sensitive to LysC enzyme) is present just upstream of the expanded CGG repeats while the remaining sequence, notably the ACG codon is untouched. Proteomic analysis of FMRpolyG after LysC digestion confirmed that FMRpolyG translation initiates at the ACG near-cognate codon (supplemental figure 1D). Of interest, proteomic analysis also revealed that the initial amino acid of FMRpolyG is a methionine suggesting that the ACG

codon is decoded by an initiator Met-tRNA despite imperfect match (Figure 1C and supplemental figure 1D). This ACG near-cognate codon is embedded in a potential Kozac consensus sequence with a purine and a guanine in -3 and +4 positions, and is conserved among multiple species (supplemental figure 1E). Translation initiation at the ACG near-cognate codon is predicted to result in a small FMRpolyG protein composed of a short 12 amino acids N-terminus, a central glycine stretch which length corresponds to the number of expanded CGG repeats and a C-terminus of 42 amino acids with no predicted structure or homology (Figure 1D). Thus, according to the number of CGG repeats, FMRpolyG may range from 6 kDa in control individuals with 30 CGG repeats to ~15 kDa in carriers of premutation with 150 CGG repeats.

A minimum of 60 to 70 expanded CGG repeats is required to detect FMRpolyG.

In contrast to RAN translation that initiates within the expanded repeats (Cleary and Ranum 2014), initiation of FMRpolyG occurs at a near-cognate codon located upstream of the repeats. This result questions whether FMRpolyG is encoded by a classic open reading frame, and if so what are the consequences of CGG expansion for FMRpolyG expression. Upstream ORFs (uORF) are short open reading frames that are located in the 5'UTR of mRNAs upstream to the main ORF. Due to the 5' to 3' scanning ribosome process and in absence of ribosome reinitiation, the translation of a uORF generally impairs translation of the downstream main ORF. Thus, to avoid complete translation inhibition, uORFs typically start by an AUG or a near cognate codon (GUG, CUG, UUG, ACG, etc.) embedded in a poor Kozac consensus sequence that enables leaky ribosomal scanning, hence translation initiation at the downstream main ORF (review in Sonenberg and Hinnebusch, 2009). To test the presence of a putative uORF in *FMR1*, we cloned various lengths of expanded CGG repeats within the human 5'UTR of *FMR1* fused in the glycine frame with a small FLAG tag (8 amino acids, ~1 kDa). Immunoblotting revealed that expression of FMRpolyG-FLAG is detected only with expanded CGG repeats over 60 to 70 CGG repeats (Figure 2A). This size of expansion is strikingly the threshold over which premutation carriers are at risk of developing FXTAS (Jacquemont et al., 2006; Apartis et al., 2012; Tassone et al., 2012), suggesting that FXTAS patients may express FMRpolyG while controls might not. Fusion of FMRpolyG in the glycine frame to a larger tag, the GFP (25 kDa), confirmed that translation occurred with expanded CGG repeats of various lengths (70 to 100 repeats) characteristic of premutation carriers (Figure 2B). However, FMRpolyG-GFP was also translated with short stretches of CGG repeats (30) found in control individuals, or even without any CGG repeats

(Figure 2B). These results suggest that translation initiation occurs at the ACG near cognate codon independently of the CGG expansion, supporting that FMRpolyG is encoded by a classic ORF. Furthermore, detection of FMRpolyG fused to a small tag only with expansion over 70 CGG repeats or with any repeat lengths when fused to a sufficiently large tag (GFP) suggest that the size of the CGG expansion is crucial for detection of FMRpolyG but not for its translation. This is characteristic of short upstream ORFs that are generally translated into small and most often undetectable peptides, but which are detected when fused with large tags, resulting in stable and detectable proteins (Aspden et al., 2014).

Next, we noted that FMRpolyG and FMRP ORFs are in different frames with the last 20 amino acids of FMRpolyG naturally overlapping the N-terminal part of FMRP (Figure 1D). Thus, translation of the FMRpolyG uORF may potentially impair ribosomal re-initiation to the downstream FMRP ORF. To test that hypothesis, we fused the 5'UTR of *FMR1* to a FLAG tag in the glycine frame and also fused the downstream FMRP ORF in frame to a GFP tag (Figure 2C). This construct expresses the FLAG-tagged FMRpolyG uORF or the downstream GFP-tagged FMRP main ORF (Figure 2C), suggesting that translation initiation occurs indeed at the FMRpolyG ACG near-cognate codon or at the FMRP ATG codon. Deletion of the 5'UTR sequence containing the ACG near-cognate initiation codon abolished expression of the FLAG-tagged FMRpolyG uORF, but enhanced translation of the downstream GFP-tagged FMRP ORF in mammalian cells (Figure 2C). In contrast, mutation of the ACG near-cognate codon into a cognate AUG initiation codon predictably enhanced translation of the FLAG-tagged FMRpolyG uORF, but abolished expression of the downstream GFP-tagged FMRP ORF (Figure 2C). Overall, these results suggest that the 5'UTR of human *FMR1* may contain an upstream ORF encoding FMRpolyG, which overlaps with the downstream FMRP ORF and reduces its translation.

To confirm that FMRpolyG is encoded by an ORF initiating upstream of the CGG repeats, we developed monoclonal mouse antibodies directed against the 12 amino acids located upstream of the glycine repeats and which thus constitute the N-terminal part of FMRpolyG (supplemental figure 2A). Immunofluorescence revealed presence of FMRpolyG N-terminus in brain sections of FXTAS patients but not in age-matched control individuals (Figure 2D). FMRpolyG N-terminus is detected as single nuclear inclusions that co-localize with ubiquitin, which is consistent with the known histopathological features of FXTAS (Greco et al., 2002). We confirmed these results by immunoblotting and found that according to the size of the

CGG expansion, the FMRpolyG protein was detected as a 10 to 14 kDa protein in the insoluble fraction of brain lysate of FXTAS individuals (Figure 2E). Importantly, FMRpolyG protein was not detected in brain samples of age-matched control individuals (Figure 2E). The poor quality of the immunoblotting is probably due to the tendency of this glycine-rich protein to aggregate. To insure that translation successfully passes across the expanded CGG repeats, we also developed a monoclonal mouse antibody against the amino acids located downstream of the glycine repeats, thus recognizing the C-terminal part of the FMRpolyG protein (supplemental figure 2A). Immunofluorescence confirmed presence of FMRpolyG C-terminus in brain sections of FXTAS patients but not in age-matched control individuals (supplemental figure 2B), which is fully consistent with previous analyses (Todd et al., 2013; Buijsen et al., 2014). Finally, translation of expanded CGG repeats into a polyalanine-containing protein was reported in transfected cell models of FXTAS (Todd et al., 2013; Kearse et al., 2016). We developed monoclonal antibodies against the C-terminal part of this protein (supplemental figure 2C). However, we observed no or very little FMRpolyA in individuals with FXTAS (supplemental figure 2D), suggesting that expanded CGG repeats are mainly translated in the glycine frame. Overall, these results confirm that expanded CGG repeats belong to an ORF embedded within the 5'UTR of *FMR1*, which translation starts upstream of the CGG repeats and is detected mainly with expansion over 70 CGG repeats resulting in a small protein ranging from ~10 kDa with 70 CGG repeats to ~15 kDa with 150 CGG repeats.

Translation of expanded CGG repeats into FMRpolyG is pathogenic in mice.

The presence of both CGG RNA aggregates, which can titrate out RNA binding proteins, and FMRpolyG–positive inclusions in individuals with FXTAS questions which pathogenic mechanism is driving neuronal degeneration in this disease. To differentiate between these two hypotheses, we developed two transgenic mouse models. The first one contains the full human 5'UTR of *FMR1* with expanded 99 CGG repeats that shall express both CGG RNA and FMRpolyG protein, while the second mouse model also expresses 99 CGG repeats but the non-canonical ACG initiation codon and surrounding 5'UTR sequence is deleted, such that it only expresses the CGG RNA (Figure 3A). Both constructs are driven by the strong chimeric ubiquitous CAG promoter, and inserted by homologous recombination within the neutral *Rosa26* mouse locus to avoid any bias due to random insertion of the transgenes. We confirmed by southern blot and PCR the correct insertion of the transgenes into the *Rosa26* locus, the absence of concatamerization at the locus and the presence of 99 CGG repeats,

which were stably transmitted with no obvious contraction or expansion, at least in the few generations (~8) of this study. To control expression of the transgenes, three upstream SV40-polyadeneylation sites bordered by loxP sites limit transcription of the expanded CGG repeats. Hence, expression and potential pathogenicity of the expanded CGG repeats is permitted only in offspring of the transgenic CGG mice crossed with mice expressing Cre recombinase (Figure 3A).

Deletion of the loxP cassette using a ubiquitously and embryonically expressed Cre recombinase (Birling et al., 2012) led to high expression of RNA with expanded CGG repeats throughout the brain, heart and liver, with lower expression in skeletal muscle, kidneys and other organs (Figure 3B). Importantly, transgene RNA expression was similar between bigenic CMV-cre mice with the full or mutant *FMR1* 5'UTR (Figure 3B). However, we found no or only very rare CGG RNA foci in brain sections of full or mutant *FMR1* 5'UTR transgenic mice at any age analyzed (supplemental figures 3A and 3B). This is consistent with the rare occurrence of CGG RNA aggregates in other transgenic mice model expressing expanded CGG repeats (Sellier et al., 2010).

Concerning the FMRpolyG protein, immunohistochemistry assays using an antibody directed against the N-terminus of FMRpolyG demonstrated expression and accumulation of nuclear aggregates of FMRpolyG in brain sections from the full 5'UTR *FMR1* transgenic mice, but not in the mutant 5'UTR mice (Figure 3C). Similar results were obtained with an antibody targeting the C-terminus of FMRpolyG (supplemental figure 3C). As observed in brain samples of individuals with FXTAS, aggregates of FMRpolyG co-localized with ubiquitinated inclusions (Figure 3D). Nuclear aggregates of FMRpolyG accumulated over time, with the largest burden of inclusions occurring within the hypothalamus, mirroring transgene mRNA expression (supplemental figures 3C and 3D). Consistent with the widespread expression of the full 5'UTR *FMR1*, some rare aggregates of FMRpolyG were found in other tissues than brain. This is reminiscent of observation of rare FMRpolyG aggregates in non-CNS tissues in FXTAS patients (Buijsen et al., 2014). In contrast, we did not observe aggregates of FMRpolyA in the full 5'UTR *FMR1* transgenic mice or in the mutant 5'UTR mice (supplemental figure 3E). These results confirm in mouse models that translation of the CGG repeats starts upstream of the expansion in the glycine frame.

To determine the consequences of FMRpolyG production *in vivo* in mice, we conducted a

series of behavioral and locomotor assays on both mouse lines. Mice with the full 5'UTR of *FMR1* develop obesity at 6 months of age. Therefore, behavioral tests were performed at 3 months of age when weight is identical between full and mutant *FMR1* 5'UTR transgenic mice. Importantly, we observed that only mice with the full 5'UTR sequence of *FMR1* and expressing the FMRpolyG protein present locomotor deficiency (supplemental videos 1 and 2), with increased falling from rotarod (Figure 3E), decreased ability of traction from the hind limbs (Figure 3F), decreased grip strength (Figure 3G) and decreased number of rears in open field observation (Figure 3H). At six months of age, mice with the full 5'UTR of *FMR1* loss mobility and develop obesity, while mice with the mutant 5'UTR that express only the CGG RNA remain normal (Figures 3I). We did not observe massive neuronal cell death, but we found some loss of Purkinje cells in mice with the full 5'UTR of *FMR1* compared to control or mutant 5'UTR mice (Figure 3J and supplemental figure 3F). Furthermore, Iba1 and Gfap staining were mildly increased in brain sections of the full 5'UTR of *FMR1* compared to mutant 5'UTR mice or control non-transgenic animals, suggesting that neuronal alterations and neuroinflammation were increased in FMRpolyG expressing animals (supplemental figure 3G). Finally, the survival curve indicates that expression of FMRpolyG is deleterious since mice expressing the full 5'UTR of *FMR1* die around 10 months, while mutant *FMR1* 5'UTR mice exhibit normal longevity and are indistinguishable from control mice (Figure 3K). These results indicate that translation of expanded CGG repeats into FMRpolyG is pathogenic, while expression of mutant RNA with expanded CGG repeats in isolation is insufficient to elicit phenotypes in mice over the time course (15 months) of this study.

To assess the tissue origin of these phenotypes, we next analyzed offspring of full 5'UTR transgenic mice crossed with Nestin-cre mice, which express the Cre recombinase in precursors of neurons and glia cells around E10.5 (Isaka et al., 1999)(Figure 4A). Quantitative RT-PCR revealed high expression of expanded CGG RNA only in mouse brain of the double transgenic Nestin-cre/full 5'UTR *FMR1* mice, confirming the specificity of the Nestin promoter (Figures 4B). Immunohistochemistry using antibodies directed against the N- or the C-terminal parts of FMRpolyG indicated expression and accumulation of nuclear aggregates of FMRpolyG in brain sections of these bigenic mice (Figure 4C). Consistent with the brain-restricted expression of the full 5'UTR *FMR1* transgene, no FMRpolyG aggregates were found in tissues other than brain. Importantly, locomotor testing at 3 months of age revealed increased falling from rotarod compared to control non-transgenic mice (Figure 4D). Histopathological analyzes revealed some Purkinje cell loss (Figure 4E) and sign of

neuroinflammation (Figure 4E) in 9 to 10 months old Nestin-cre/full 5'UTR *FMR1* mice compare to control animals. Interestingly, these mice expressing FMRpolyG only in the brain developed obesity (Figure 4G) and present a reduced longevity (Figure 4H) similarly to the CMV-cre/full 5'UTR *FMR1* mice, demonstrating that these phenotypes are of neuronal origin.

The 5'UTR sequence of *FMR1* impairs formation of expanded CGG repeats RNA foci.

The absence of RNA foci of expanded CGG repeats in both full and mutant 5'UTR *FMR1* mouse models is puzzling as it was proposed that FXTAS results from an RNA gain-of-function mechanism in which the nuclear accumulation of mutant CGG RNA would sequester various RNA binding proteins. Thus, we tested whether transfection of the 5'UTR *FMR1* with expanded CGG repeats construct would form any RNA foci in mouse neuronal cell cultures (supplemental figure 4A). As positive control, we used a construct expressing expanded CGG repeats but deleted of any *FMR1* sequence and that was shown to form RNA foci upon transfection in cells (Sellier et al., 2010; Sellier et al., 2013) (supplemental figure 4A). RNA FISH assays indicated that expanded CGG repeats embedded in the 5'UTR of *FMR1* formed rare RNA foci compared to the expanded CGG repeats deleted of *FMR1* sequence (supplemental figure 4B). Furthermore, RT-PCR performed on nuclear and cytoplasmic fractions indicated that most of RNAs containing expanded CGG repeats embedded within the 5'UTR of *FMR1* were exported from the nucleus to the cytoplasm (supplemental figure 4C). In contrast, expanded CGG repeats RNA without *FMR1* sequence is mainly retained within cell nuclei (supplemental figure 4C). These results highlight the nuclear retention bias induced by using artificial constructs in which microsatellite repeats are separated from their natural sequence context. Importantly, these data also indicate that CGG expanded repeats embedded in their natural *FMR1* sequence are exported into the cytoplasm and thus available for translation into the FMRpolyG protein.

The C-terminal part of FMRpolyG contributes to its toxicity.

As translation of the CGG expanded repeats appears to drive pathogenicity in FXTAS animal models, we next investigated by which mechanisms the FMRpolyG protein may elicit neuronal cell dysfunction. Immunofluorescence of FMRpolyG in primary cultures of embryonic mouse cortical neurons transfected with 99 CGG repeats embedded within the natural human 5'UTR of *FMR1* fused to GFP in the glycine frame indicated that FMRpolyG first accumulates in the cytoplasm where it forms aggregates that migrated within the cell

nucleus (Figure 5A). To exclude any bias caused by the GFP tag, we repeated that experiment with 99 CGG repeats embedded in the 5'UTR of *FMR1* fused to a FLAG tag in the glycine frame and observed identical results in primary mouse cortical neurons (supplemental figure 5A), as well as in HEK293 cells (supplemental figure 5B). Furthermore, immunoblotting on the soluble and insoluble fractions of neuronal cells transfected with 99 CGG repeats embedded within the natural human 5'UTR of *FMR1* fused to GFP in the glycine frame indicated that FMRpolyG progressively accumulates in the insoluble fraction, which is consistent with its propensity to form aggregate (supplemental figure 5C).

To identify the sequence driving FMRpolyG aggregation, we cloned various deletion mutants of FMRpolyG expressing either its N-terminus with the glycine repeats or its C-terminus in isolation. As a control, expression of the full-length FMRpolyG protein in primary cultures of E18 mouse cortical neurons leads to nuclear aggregates associated with cell death (Figure 5B). Expression of the polyglycine stretch in isolation, deleted of FMRpolyG C-terminus, was sufficient to elicit aggregation. However, these aggregates were not associated with a significant increase in neuronal cell death above control non-transfected or GFP-expressing neurons. In contrast, expression of GFP fused to the 42 amino acids constituting the C-terminus of FMRpolyG caused neuronal cell death without forming nuclear aggregates (Figure 5B and supplemental videos 3 to 6). We confirmed these results in a second cell model. Transfection of neuronal Neuro2A cells demonstrated that expression of the polyglycine stretch in isolation was not overly toxic, while expression of the C-terminal part of FMRpolyG induced cell death (supplemental figure 5D). As further control and to exclude any bias due to the GFP tag, we fused the FMRpolyG constructs to a smaller FLAG tag. Expression of FMRpolyG-FLAG was toxic as it reduce neuronal cell viability by half, while expression of the polyglycine stretch deleted from FMRpolyG C-terminus did not induce major cell death (supplemental figure 5E).

To confirm these results in an animal model and test the toxicity of FMRpolyG on a longer time period, we developed *Drosophila* transgenic lines expressing either the full FMRpolyG protein or the polyglycine stretch in isolation under a *UAS* promoter. Toxicity was assessed by two separate assays. First, *UAS* FMRpolyG-GFP and *UAS* polyG ΔCter-GFP flies, both expressed under an ATG initiator codon, were crossed with an *Act5c-Gal4* driver line, which leads to ubiquitous expression of the transgene during development. Total progeny carrying either the transgenes or a balancer chromosome were then quantified over 3 independent

crosses. Importantly, expression of the full FMRpolyG was toxic and reduced progeny eclosion by half, while the eclosion rate was only slightly reduced in flies expressing the polyglycine stretch in isolation (Figure 5C). To exclude any potential bias due to random insertion effects, we generated and analyzed various independent lines expressing either FMRpolyG or its polyglycine stretch. As control, quantitative RT-PCR showed similar expression of the FMRpolyG or polyglycine transgenes (supplemental figure 5F). Importantly, all lines expressing FMRpolyG-GFP showed reduced viability compared to GFP controls. In contrast, the lines expressing the glycine repeats in isolation consistently presented milder phenotype (supplemental figure 5G).

As a second measure of FMRpolyG toxicity, we crossed these same transgenic fly lines to a *Tubulin-Gal4* Geneswitch driver, which expression is induced upon addition of mifepristone (RU-486). As control, adult transgenic flies reared off of RU-486 drug exhibited no differences in viability compared to control flies, indicating that insertion of the FMRpolyG or polyglycine transgenes had no deleterious effect. In contrast, transgenic flies feed with RU-486, activating ubiquitous transgene expression, showed a decrease in viability over time for FMRpolyG expressing flies (Figure 5D). Interestingly, adult *Drosophila* expressing the polyglycine stretch alone also presented a mild decrease of their viability (Figure 5D). The difference of toxicity of the polyglycine stretch between cell culture and *Drosophila* is probably due to the longer time of expression in the latter, revealing a mild toxicity undetected in the short time frame of cell transfection. Overall, these results demonstrate that expression of FMRpolyG is pathogenic, with its polyglycine stretch driving aggregation and its C-terminus enhancing toxicity in neuronal cells as well as in animal models.

FMRpolyG interacts with LAP2ß and alters the nuclear lamina.

Next, we reasoned that if the C-terminal region of FMRpolyG was contributing to its toxicity, this might be through interaction with other proteins. To identify FMRpolyG binding proteins, we performed a tandem tag purification of HA-FLAG tagged FMRpolyG transfected into Neuro2A cells followed by nano-LC-MS/MS analysis of associated proteins. This approach identified various FMRpolyG-associated proteins (supplemental table 1), among which the most prominent was Lap2ß (Figure 6A). Lamina associated polypeptide 2 (LAP) alpha and beta are two isoforms of the LAP2 protein that differ in their C-termini, which originate from alternative splicing of the *TMPO* pre-mRNA. LAP2α is diffusely localized in the nucleus, while LAP2ß carries a transmembrane domain in its C-terminus that anchors it to the inner

nuclear membrane (Foisner and Gerace 1993; Furukawa et al., 1995). LAP2ß interacts with lamin B1 and B2 and help to organize these proteins near the nuclear inner membrane. Consequently, alteration of LAP2ß results in disorganization of the nuclear lamina architecture (Dubinska et al., 2015; Gant et al., 1999). Since alteration of the nuclear lamina has been observed in FXTAS (Arocena et al., 2005; Iwahashi et al., 2006; Hoem et al., 2011), we pursued the study of LAP2 further.

Of interest, a HA-FLAG tagged construct containing the polyglycine stretch deleted of FMRpolyG C-terminal part did not pull down Lap2ß (Figure 6A). Co-immunoprecipitation studies confirmed the association of LAP2ß with HA-FLAG tagged FMRpolyG, but not with the polyglycine stretch in isolation (Figure 6B). To exclude a potential bias of the HA-FLAG double tag, we repeated that experiment with FMRpolyG fused to the GFP. Co-immunoprecipitation assays confirmed that interaction with LAP2ß required the C-terminus of FMRpolyG as its polyglycine stretch alone was not co-immunprecipitated by HA-tagged LAP2ß (Figure 6C). Inversely, a construct expressing only the C-terminal part of FMRpolyG interacted with LAP2ß (Figure 6C).

An interaction between LAP2ß and FMRpolyG questions whether translation of the expanded CGG repeats in FXTAS may alter the localization or the function of LAP2ß. Expression in HEK293 cells of 99 CGG repeats embedded within the 5'UTR of *FMR1* fused to the GFP in the glycine frame resulted in formation of nuclear aggregates of LAP2ß that co-localized with FMRpolyG-GFP inclusions (supplemental figure 6A). Identical results were obtained with CGG repeats tagged in the glycine frame with a small FLAG tag (supplemental figure 6A). Similarly, expression of FMRpolyG-GFP in primary cultures of cortical neurons from mouse embryo indicated that FMRpolyG recruits endogenous Lap2ß in nuclear aggregates (Figure 6D). Interestingly, a construct containing only the polyglycine stretch of FMRpolyG did not alter Lap2ß localization, suggesting that the C-terminal part of FMRpolyG is important for interaction and delocalization of LAP2ß (Figure 6D). We confirmed these results in brain sections of mouse models of FXTAS. Immunohistochemistry assays indicated that Lap2ß accumulates in nuclear aggregates in brain sections of the full 5'UTR *FMR1* transgenic mice (Figure 6E). In contrast, Lap2ß was normally localized at the nuclear lamina in non-transgenic animals or in the mutant 5'UTR *FMR1* transgenic mice (Figure 6E). As observed in transfected cells, aggregates of Lap2ß co-localized with FMRpolyG inclusions in mouse brain sections (Figure 6F). We next tried to confirm these results in brain sections of

individuals with FXTAS. However, LAP2ß immunohistochemistry were of poor quality on human autopsied material. We nevertheless observed some LAP2ß aggregates in cerebellum sections as well as in the hippocampal area of FXTAS individuals (Figure 6G and supplemental figure 6B). These aggregates of LAP2ß were absent in the cerebellum from age-matched control cases (Figure 6G). Furthermore, immunofluorescence assays indicated that LAP2ß co-localized with FMRpolyG inclusions in individuals with FXTAS but not in age-matched control individuals (Figure 6H).

Alteration of LAP2ß results in disorganization of the nuclear lamina (Gant et al., 1999; Dubinska et al., 2015). Thus, we next investigated the consequences of FMRpolyG expression on the architecture of the nuclear lamina. Expression of 99 CGG repeats embedded within the 5'UTR of *FMR1* fused in the glycine frame with either a GFP or a FLAG tag in HEK293 cells resulted in disorganization of the nuclear lamina as evidenced by alterations of the lamin B1 labeling (supplemental figure 6C). Similarly, expression of FMRpolyG-GFP in primary cultures of mouse embryonic cortical neurons altered lamin B1 nuclear organization (supplemental figure 6D). In contrast, expression of the polyglycine stretch in isolation did not affect nuclear lamina organization (supplemental figure 6D). Furthermore and as reported previously with lamin A (Iwahashi et al., 2006), the localization of lamin B1 was altered in brain sections of FXTAS compared to controls (supplemental figure 6E).

LAP2ß and nuclear lamina are altered in neuron differentiated from FXTAS iPS cells.

Alterations of the nuclear lamina were observed in autopsied samples that may represent an end stage of the disease. Thus to overcome this potential bias, we next investigated organization of the nuclear lamina in neuronal cells expressing endogenous levels of FMRpolyG. To accomplish this, we developed human induced pluripotent stem (iPS) cells derived from fibroblasts from two age-matched controls and three different FXTAS patients with expansion of 84, 90 and 99 CGG repeats. Fibroblasts were successfully reprogrammed using retroviruses expressing Oct4, Sox2, Nanog and Lin28 (Jung et al., 2014). In addition to karyotype analyses, which were normal for all cell lines, retroviral silencing as well as stemness and pluripotency were confirmed by classic RT-qPCR and teratoma assays. Of interest, expanded CGG repeats were stable with no contraction or expansion in iPS clones compared to fibroblasts. Since FXTAS affect various brain regions, including the cortex, iPS cells from control and FXTAS individuals were differentiated into homogenous populations of telencephalic neurons (Marteyn et al., 2011; Boissart et al., 2013). Number of CGG repeats

was stable during differentiation, and we observed no gross alterations or delay in neuronal differentiation of FXTAS iPS cells compared to controls (supplemental figure 7A). As observed previously (Liu et al., 2012), expression of *FMR1* mRNA was two to three fold increased in neurons derived from FXTAS iPS cells compared to controls (supplemental figure 7B). This is consistent with the increased levels of *FMR1* mRNA observed in carriers of a CGG premutation (Tassone et al., 2000; Kenneson et al., 2001; Tassone et al., 2007).

Immunofluorescence using the antibody against the N-terminal part of the FMRpolyG protein detected accumulation of nuclear aggregates of FMRpolyG in FXTAS neurons, but not in control neurons (Figure 7A). FMRpolyG aggregates accumulated over time post-differentiation, with 5 to 10% of neurons exhibiting small FMRpolyG aggregates at 20 days of differentiation, while 20 to 30% of neurons present FMRpolyG nuclear aggregates after 40 days of differentiation (supplemental figure 7C). In contrast, RNA foci of expanded CGG repeats were rare or absent in FXTAS neurons at 40 days of differentiation (supplemental figure 7C). These results confirm observations in transgenic mice and indicate that translation of expanded CGG repeats into FMRpolyG is readily observed while formation of RNA foci of expanded CGG repeats is rare in animal and iPS cell models of FXTAS, at least in the time frame of our study.

Importantly, immunofluorescence analyzes revealed that endogenous LAP2ß loss its normal localization and forms nuclear inclusions that co-localize with FMRpolyG in neuron differentiated from FXTAS iPS cells (Figure 7A). Furthermore, aggregation of LAP2ß and FMRpolyG were associated with disorganization of the nuclear lamina structure as shown by alteration of the lamin B1 labeling in FXTAS neurons (Figure 7B). In contrast, neurons differentiated from iPS cells of control individuals exhibited normal LAP2ß and lamin B1 localization (Figures 7A and 7B). These results fully confirm observation in transgenic mice and brain sections of FXTAS individuals and demonstrate that alteration of the nuclear lamina in FXTAS is unlikely a bias of autopsied material but a pathogenic consequence of the translation of expanded CGG repeats into FMRpolyG.

Overexpression of LAP2 rescues neuronal cell death caused by FMRpolyG.

If binding of FMRpolyG to LAP2ß is a key pathogenic event in FXTAS, then resupplying excess LAP2ß may suppress FMRpolyG-induced toxicity. To address this question, we co-expressed LAP2ß or a control vector expressing the RFP with FMRpolyG in mouse neuronal

cultures. LAP2ß co-expression was sufficient to rescue the cell death induced by FMRpolyG-GFP co-expression (Figure 5A). Consistent with LAP2ß binding to the C-terminal part of the FMRpolyG protein, expression of LAP2ß also rescued cell death caused by expression of the FMRpolyG C-terminus in isolation (Figure 7C). As controls, expression of LAP2ß had no effect on neuronal survival in control GFP-transfected cells or neurons expressing the polyglycine fragment in isolation. As further controls, we repeated that experiment with a construct expressing FLAG-tagged FMRpolyG to exclude any bias from the GFP tag. Expression of LAP2ß rescued cell toxicity caused by FMRpolyG-FLAG (supplemental figure 7D), confirming that the interaction FMRpolyG with LAP2ß is an important pathogenic event in FXTAS. Overall, these data suggest a mechanism by which FMRpolyG can elicit toxicity (Figure 8), and an explanation for the previously observed nuclear lamina disorganization in FXTAS (Arocena et al., 2005; Iwahashi et al., 2006; Hoem et al., 2011).

DISCUSSION

Previous pioneering studies have demonstrated that FXTAS is caused by expression of mutant RNAs containing expanded CGG repeats (Willemsen et al., 2003; Jin et al., 2003; Arocena et al., 2005; Hashem et al., 2009; Entezam et al., 2007; Hukema et al., 2014; Hukema et al., 2015). However, whether expanded CGG repeats are pathogenic through an RNA gain-of-function mechanism or through translation into a toxic protein was unclear. Using novel mouse models expressing or not FMRpolyG, our study now demonstrates a direct role for FMRpolyG in CGG repeat associated toxicity. In contrast, expression of expanded CGG repeats as RNA in isolation is not sufficient to induce pathogenicity. This study fully confirms into a mammalian system previous results obtained in *Drosophila* (Todd et al, 2013) and establishes that FMRpolyG synthesis is required for formation of ubiquitin-positive inclusions in mouse. Moreover, we found that FMRpolyG disrupts nuclear lamina architecture through binding to the LAP2ß protein. In some aspect this is reminiscent to Amyotrophic Lateral Sclerosis and FrontoTemporal Dementia (ALS-FTD) in which expanded GGGGCC repeats in the C9ORF72 gene (Renton et al., 2012; De-Jesus et al., 2012) are RAN translated into pathogenic di-peptide containing proteins (Ash et al., 2013; Mori et al., 2013; Zu et al., 2013) that disrupt nucleocytoplasmic transport proteins (Zhang et al., 2015; Freibaum et al., 2016; Jovičić et al., 2015; Zhang et al., 2016).

However, in contrast to the RAN translation of microsatellite repeats that initiates within the expanded repeats (Cleary and Ranum 2014), our work confirm that expression of FMRpolyG depends of an initiation to a near-cognate codon located upstream of the CGG repeats (Todd et al., 2013; Kearse et al., 2016). Moreover, we provide mass-spectrometry data that the N-terminal amino acid of FMRpolyG is a methionine in mammalian cells. Finally, our data suggest that expanded CGG repeats belong to a potential small ORF translated upstream to the main FMRP ORF. We noted that translation of the FMRpolyG uORF reduces the expression of FMRP, thus expression of FMRpolyG may contribute to the slight decrease of FMRP expression observed in FXTAS individuals (Tassone et al., 2000; Kenneson et al., 2001; Tassone et al., 2007). However, other mechanisms may also contribute to this decrease expression of FMRP in FXTAS, including the increase ribosome stalling observed with expanded CGG RNA hairpin structure (Feng et al., 1995; Primerano et al., 2002). Ribosome profiling and bioinformatics analyses reveal that upstream ORF (uORF) are commons in mammalian mRNAs and can initiate at non-canonical codons (Calvo et al., 2009; Ingolia et al., 2011; Fritsch et al., 2012; Ji et al., 2015). Mutations in uORF are known to cause human

diseases, but mostly if not exclusively, by altering the expression of their downstream ORF (review in Barbosa et al., 2013). In contrast, we propose here that FXTAS is characterized by a mutation extending the length of a uORF resulting in expression of a toxic polyglycine-containing protein, FMRpolyG. This model is reminiscent of expansion of tri-nucleotides repeats in ORF resulting in expression of pathogenic polyglutamine or polyalanine-containing proteins.

Interestingly, this translation model may provide some molecular basis to the threshold of severity and incomplete penetrance observed in FXTAS. Indeed, while the prevalence of carriers of a premutation, which is defined as an expansion between 55 and 200 CGG repeats, ranges from 1:130 to 1:256 for females and from 1:250 to 1:813 for males according to ethnic groups and countries (review in Tassone et al., 2012), it is estimated that only 1 in 3,000 men older than 50 years will develop FXTAS. This less than expected frequency is explained by an incomplete penetrance but also by most of FXTAS cases occurring in the range of ~70 to ~150 CGG repeats, with exceptional FXTAS cases below 70 CGG repeats where most (80%) of premutation alleles are found (Jacquemont et al., 2006; Apartis et al., 2012; Tassone et al., 2012). Importantly, our experiments demonstrate that expression of small-tagged FMRpolyG is hardly detectable below 70 CGG repeats. This is consistent with the difficulties to detect small proteins below 10 kDa, but also with the increase expression of mutant CGG RNA in carriers of premutation above 70 CGG repeats (Tassone et al., 2000; Kenneson et al., 2001; Tassone et al., 2007), as well as with the proposed increase stalling of the ribosome to the CGG hairpin structure, which would promote translation initiation to near-cognate codons located upstream to the CGG repeats (Todd et al., 2013). Furthermore, previous work demonstrated that translation is impaired by expansion of CGG above 150 CGG repeats (Feng et al., 1995; Primerano et al., 2002). Thus, we propose that most of FXTAS cases may be restricted to a more limited premutation range (70-150) of CGG expansion since expression of the toxic FMRpolyG protein is reduced below ~70 and above ~150 CGG repeats. Interestingly, this pathogenic range (70-150) of CGG repeats correlates with presence of nuclear ubiquitin-positive aggregates in brain sections of a mouse model of FXTAS (Brouwer et al., 2008), as well as in individuals with FXTAS (Greco et al., 2006). Thus, we propose that productive translation of CGG repeats into FMRpolyG between 70 and 150 CGG repeats may explain the lesser incidence of FXTAS cases observed, compared to the relatively high frequency of CGG premutations that are mostly below 70 repeats.

Of interest, a pathogenic mechanism where expansion of nucleotide repeats into a uORF results in expression of a toxic protein may apply to other diseases, such as Fragile X-associated Primary Ovarian Insufficiency (FXPOI) caused alike FXTAS by expanded CGG repeats in the *FMR1* gene (Buijsen et al., 2016). Similarly, it is striking to note that expanded GGGGCC repeats, which are located upstream to the C9ORF72 ORF and are the main genetic cause of ALS-FTD, are in frame with an upstream CTG near-cognate codon in a correct Kozac sequence (gctCTGg) to encode a Glycine Alanine-polypeptide, which is the most common dipeptide-repeats protein detected in individuals with ALS-FTD (Mackenzie et al., 2015). While hypothetical, extending further the FXTAS/ Fragile X model to C9ORF72 might predict that long expansions of GGGGCC repeats in tissues prone to somatic expansions would be transcriptionally silent due to epigenetic modification of the promoter, while shorter expansions in tissues prone to repeats contraction would be transcribed and translated.

One of the main conclusions of this work is that mice expressing FMRpolyG develop a severe phenotype with reduced longevity, while mice expressing only the expanded CGG repeats RNA are indistinguishable from control mice. These results suggest that accumulation of RNA with expanded CGG repeats or potential RAN translation initiating within these repeats is not overly pathogenic in mice, at least in the time frame of our study. While informative on the pathogenicity of FMRpolyG, these mouse models present some limitations. Notably, our behavioural and locomotor investigations were limited to 3 months aged mice since animals expressing FMRpolyG in the brain develop obesity and grew immobile around 6 months of age. Of interest, the obesity developed by the FMRpolyG expressing mice is probably caused by dysfunction of neurons from the hypothalamus, which regulate rodent feeding behavior, as this brain area is the one expressing the most and the earliest FMRpolyG. Moreover, we found alteration of the expression of the leptin receptor by RT-qPCR in the hypothalamic area of these mice. This is in contrast to individual with FXTAS who do not develop morbid obesity and in whom little ubiquitin-positive aggregates are observed into the hypothalamus (Greco et al., 2002). Furthermore, no FMRpolyG-expressing mice survived past 1 year of age, impairing any chance to observe potential neurodegeneration at later age. While the cause of death in these mice remains to be carefully investigated, we believe that it is of neuronal origin as mice expressing FMRpolyG under a Nestin-cre driver present similar motor impairments, obesity and shortened survival. Thus and as reported previously (Hashem et al.,

2009), analysis of potential Purkinje cell dropout and ataxia at old age will require specific expression of FMRpolyG in mouse cerebellum.

Comparison of these novel mice with previous models of FXTAS expressing either mRNA with 90 CGG repeats and FMRpolyG protein (Willemsen et al., 2003; Hashem et al., 2009; Hukema et al., 2015) or only mRNA with ~118 CGG repeats (Entezam et al., 2007; Qin et al., 2011) is difficult as these pioneering mice express either endogenous level of transgene due to knockin of the expanded CGG repeats within the endogenous *Fmr1* gene (Willemsen et al., 2003; Entezam et al., 2007), or express the expanded repeats in more restricted regions of the brain (Hashem et al., 2009; Hukema et al., 2015) compared to the mouse models developed in the present study. It is noteworthy that knockin mice with expanded CGG repeats, but that do not express FMRpolyG due to presence of a stop codon located upstream of the CGG repeats (Entezam et al., 2007; Todd et al., 2013), present some Purkinje cell dropout at 2 years of age but with no ataxia or rotarod deficiency, and a mild behaviour phenotype with hyperactivity, reduced anxiety and some subtle deficits in social interaction (Entezam et al., 2007; Qin et al., 2011). These behaviour alterations are similar to the one observed in *Fmr1* knock-out mouse model, which suggest that they might be caused by the reduced expression of Fmrp observed in this CGG knockin mouse model (Qin et al., 2011). In contrast, the mild Purkinje cell dropout could be caused by expression of the mutant RNA containing expanded CGG repeats (Entezam et al., 2007). These results suggest that expanded CGG repeats are mildly toxic at the RNA level at old age. However, an important distinction not addressed previously is the timing of the pathogenic events in FXTAS models. Indeed, we observed that at time where FMRpolyG readily accumulates in nuclear inclusions, RNA foci of expanded CGG repeats are rare or undetectable both in mouse models and in neurons from human iPS cells derived from individuals with FXTAS. These data suggest a hierarchy in the pathogenic events occurring in FXTAS. Notably, the CGG RNA foci may be characteristic of an end stage of the disease, while expression of FMRpolyG may potentially represent a relevant biomarker to follow disease progression, as well as a most promising therapeutic target (Yang et al., 2015).

Finally, we found that toxicity of FMRpolyG is mediated at least in part through sequestration of the LAP2ß protein into nuclear aggregates, leading to disruption of the nuclear lamina architecture and neuronal cell death. These data provide a molecular mechanism to the previously reported nuclear lamina disorganization in FXTAS (Arocena et al., 2005; Iwahashi et al., 2006; Hoem et al., 2011). However, it is likely that the FMRpolyG protein mediates its

toxic effect through more than one mechanism and/or protein partner. Notably, LAP2ß regulates gene expression through association with various transcription factors as well as to DNA and with the DNA-binding protein BAF1 (Barrier To Autointegration Factor 1) protein (Cai et al., 2001; Shumaker et al., 2001; Nili et al., 2001; Somech et al., 2005; Naetar et al., 2008; Zullo et al., 2012). Thus, it remains to be determined whether FMRpolyG alters the transcriptional regulatory activity of LAP2. Also, proteomic analysis indicates that FMRpolyG pulls down various other proteins, including mitochondrial and proteasome proteins as well as actinin and actin related proteins. Some of these proteins are good candidates to contribute to the mitochondrial and proteasome alterations that have been observed in FXTAS (Ross-Inta et al., 2010; Kaplan et al., 2012; Hukema et al., 2014; Oh et al., 2015). Thus, further investigations are required to test the potential pathological effect of FMRpolyG on mitochondria, protein degradation mechanisms and cell cytoskeleton in FXTAS.

In conclusion, it now emerges that the *FMR1* gene encodes for two different proteins, FMRpolyG and FMRP, which levels are inversely modulated by the size of the CGG repeats expansion and which are involved in two different genetic diseases, FXTAS and Fragile X.

MATERIAL & METHODS:

Human samples.

All brain samples were obtained from the FXTAS brain repository at the UC Davis School of Medicine with the informed consent of individuals and approved by the Institutional Review Board of the University of California, Davis. Patients have been described previously (Cases LR, 58-02-WD and 1007-05-HP from Greco et al., 2006 and Pretto et al., 2014). Fibroblasts of three FXTAS male individuals with confirmed premutation of 84, 90 and 99 CGG repeats were obtained with the informed consent of individuals and approved by the Institutional Review Board of the Hospital La Pitié Salpêtrière.

Constructs.

PCMV6 containing C-terminally FLAG-tagged human cDNAs of LAP2ß was purchased from OriGene. Plasmids containing the 5'UTR of human *FMR1* fused in the glycine frame with the FLAG tag or GFP are deposited at Addgene (Plasmid #63089, #63090 and #63091). The FLAG tag is fused to the C-terminal end of FMRpolyG with a two amino acids (glycine valine) linker, and the GFP tag is fused to the C-terminus of FMRpolyG with one amino acid (glycine) linker. Mutations of the ACG into ATG or deletions of the 5'UTR of *FMR1* were achieved by oligonucleotide ligations. To insure stability of the expanded CGG repeats, all CGG plasmids were transformed into STBL3 bacterial strain (Invitrogen) and growth at room temperature (22°c).

Cell cultures, viability assays and transfections.

Neuro2A, HEK293 or HeLa cells were plated in 6 well tissue culture plates in DMEM 1 g/l glucose with 5% FCS and gentamycin. After 24 hours, cells were transfected with plasmid DNA using Lipofectamine 2000 (Invitrogen) according to manufacturer instructions. Primary cortical neurons were prepared from C57Bl/6 mice embryos of day E18 and grown on polylysine coated 24-well plates in Neurobasal Medium (NBM) supplemented with 1xB27, 0.5 mM L-glutamine and 100 IU/ml penicillin/streptomycin at 37° C with 5% CO2. Neurons were transfected at day 3 with Lipofectamine 2000 (Invitrogen) in 400 µl NBM. Medium was replaced after 3h with a 1:1 (v:v) mixture of conditioned and fresh NBM. For cell viability, cells were detached by scraping and resuspended in PBS. TO-PRO-3 iodide (Fisher scientific, T-3605) was added at 20 nM to each sample and gently mix just prior to analysis on the FACS and 30 000 cells were counted.

Generation and differentiation of iPSC into neurons.

Primary dermal fibroblasts were maintained on gelatin-coated dishes in DMEM 1 g/l glucose with antibiotic, antimycotic and 0.1 mM non-essential amino acids (Invitrogen) and 10% FBS for 5 passages. On day 1, 1x10exp5 fibroblasts were transduced by lentivirus carrying cDNAs of Oct4, Sox2, Nanog and Lin28 with 8 mg/ml of polybrene (Sigma). On day 2, medium was replaced with fresh medium and on day 3, infected cells were transferred onto a 100-mm dish containing 1x10exp6 feeder cells (passage 3 mitomycin-C treated mouse embryonic fibroblasts). From day 6 to 9, fibroblast medium was progressively switched to human induced pluripotent stem cell medium (KO-DMEM, 20% KOSR, 2 mM L-glutamine, 0.1 mM non-essential amino acids, Penicillin-Streptomycin, 0.1 mMb-mercaptoethanol supplemented with 10 ng/ml of bFGF (R&D Systems). Human iPSC clones were picked at week 4 and expanded on matrigel coated 35 mm dishes (BD Biosciences) in mTeSR1 medium (Stemcell Technologies). For embryoid body formation, hiPSC were dissociated with dispase solution (1 mg/ml, Stemcell Technologies), resuspended in 1 ml of Aggrewell medium (Stemcell Technologies) containing 2 mM Y27632 (Stemgent), centrifuged in Aggrewell plates for 3 min at 80g and further incubated at 37°C for 24 h. The next day, embryoid bodies were transferred in 3 ml of Aggrewell medium. The following days, medium was progressively switched to KO-DMEM, 20% FBS, 2 mM L-glutamine, 0.1 mM Non-Essential Amino Acids, penicillin–streptomycin and after 30 days, embryoid bodies were collected. For karyotype analysis, hiPSC cells were treated with colchicine (Sigma) for 4 h and cells were shocked with hypotonic KCl 0.075 M solution for 20 min at 37°C. Cells were fixed in methanol:acetic acid solution (3:1) and conventional cytogenetic was performed applying RHG banding on metaphase chromosomes. Expression of pluripotent surface markers of hIPSC was analyzed by FACS using anti-Tra-I-60, anti-Tra-I-81 and anti-SSEA4 antibodies (Millipore). Expression of 90 validated genes associated with stem cell pluripotency and differentiation to all three germ layers of hiPSC and corresponding embryoid bodies were analyzed using the Human Stem Cell Pluripotency Array (Applied Biosystems) according to manufacturer instructions. For in vivo teratoma formation, cells from one Matrigel coated 60 mm-dish were collected by dispase treatment and resuspended in 75 ml of KO-DMEM, mixed with 75 ml of Matrigel (BD Biosciences) and the hiPSC-Matrigel mixture was injected subcutaneously in 8-week-old NOD/SCID female mice (Charles River Laboratory, 2 mices injected for each hiPSC clone). After two to three months, teratomas were dissected and fixed in formalin, embedded in paraffin and processed with hematoxylin and eosin staining at the histology

laboratory of the Institute Clinique de la Souris (ICS). For differentiation of human pluripotent cells in Neuronal Stem cell (NSC), one B6 dish of IPS (60-80 % confluence) was washed with NFS medium (N2B27 supplemented with FGF2 5 ng/ml (Peprotech, AF-100-18B), hNoggin 260 ng/ml (R&D120-10C), SB431542 8,7 µg/ml (Tocris, 1614) containing ROCK inhibitor 3,5 µg/ml (Y-27632, Calbiochem, 688000) before to cut clumps. Clumps were collected and incubated overnight at 37°C in B3 UltraLow Attachment Dish (Corning). The next day, the clumps were transferred to a dish pre-coated with poly-ornithine 0,1% (Sigma, p4957) and laminine at 1 mg/ml (Sigma, L2020) and maintained with medium. After 24 hours, the medium was changed to NFS medium without rock inhibitor and medium was changed every two days during 8-10 days. After the appearance of neural rosette, the medium was replaced with NSC (N2B27 supplemented with FGF2 10 ng/ml, EGF 10 ng/ml (R&D, 263-EG-200RD), hBDNF 20 ng/ml (R&D, 248-BD-025CF). At confluence, cells were passaged 1 : 3 in NSC medium. For differentiation of NSC in neurons, confluent cells were dissociated with trypsin and plated on pre-coated with polyornithine (Sigma) and laminin (Sigma) in 24-well plate (50000 cells/well) in neuron medium (N2B27 supplemented with hBDNF 20 ng/ml and laminine 2 µg/ml). Media change was performed every 2 days and cells were differentiated during 55 days.

Monoclonal antibody production.

To generate monoclonal antibodies directed against FMRpolyG or FMRpolyA, 8 week old female BALB/c mice were injected intraperitoneally with KLH conjugated peptides (FMRpolyA C-ter 5FM: PRAPAAHLSGAGSRR, FMRpolyG N-ter 8FM: MEAPLPGGVRQRG or FMRpolyG C-ter 9FM: GGWASSARSPPLGGGLPALA) with 200ug of poly(I/C) as adjuvant. Three injections were performed at 2 weeks intervals and four days prior to hybridoma fusion, mice with positively reacting sera were re-injected. Spleen cells were fused with Sp2/0.Agl4 myeloma cells. Supernatants of hybridoma cultures were tested at day 10 by ELISA for cross-reaction with peptides. Positive supernatants were then tested by Immunofluorescence and western blot on transfected HeLa cells. Specific cultures were cloned twice on soft agar. Specific hybridomas were established and ascites fluid was prepared by injection of 2x106 hybridoma cells into Freund adjuvant-primed BALB/c mice. All animal experimental procedures were performed according to the French and European authority guidelines.

Western blotting.

For the small FMRpolyG-FLAG tagged or endogenous FMRpolyG (<15kDa), 20 to 50 µg of proteins were resolved by 12% bis-Tris Gel (NuPAGE) and transferred onto PVDF 0,2 µm membrane. The membrane was blocked with 5% non-fat dry milk in TBS-Tween 1% and incubated with FLAG (rabbit PA1-984B), 8FM or 9FM antibody (1:100) overnight at 4°C. Membrane was washed 3 times and incubated with secondary peroxidase antibody (1:3000, Cell Signaling) 1 hour in TBS-Tween 1%, followed by washing and ECL chemoluminescence revelation (Amersham ECL Prime). Concerning FMRpolyG-GFP tagged (>30kDa) and other large proteins analyzes, 20 µg of proteins were homogenized in 1x laemmli sample loading buffer, denatured 3 min at 95°C, separated on 4-12% bis-Tris Gel (NuPAGE), transferred on nitrocellulose membranes (Whatman Protan), blocked with 5% non-fat dry milk in TBS-Tween 1% (Tris Buffer Saline buffer), incubated with anti-FMRpolyG (8FM or 9 FM, 1/100), Lap2b (BD Biosciences 611000), GFP (Abcam ab290), GAPDH (ab125247, Abcam), HA (ThermoFisher Scientific 26183) in TBS-Tween 1%, washed 3 times and incubated with anti-rabbit or mouse Peroxidase antibody (1:3000, Cell Signaling) 1 hour in TBS-Tween 1%, followed by washing and ECL chemoluminescence revelation (Amersham ECL Prime). Concerning human brain tissue preparation, small pieces of lyophilized frozen brain tissue were homogenized in 100 µl of Tris-SDS buffer (100 mM Tris pH 9, 5 % SDS 20%, 5 % b-mercaptoethanol), boiled at 100°C during 5 min then centrifuged at 13 000 rpm for 20 min at 4 °C. The supernatant was removed. The pellet was washed twice with water and homogenized in 20 µl of formic acid and incubated at 37°C during 30 min. Next, the homogenate was dried in speed-vac and resuspended in 40 µl Laemmli loading buffer prior to western blot analysis.

Lysostaphin treatment

3x10exp5 transfected HeLa cells were scrapped in PBS 1X and centrifuged during 10 min at 3000 rpm at 4°C. The pellet was resuspended in 800 µl of RIPA. 16 µl of cell extract was incubated with 1 µg of lysostaphin (Prospec, ENZ-269) during 10 to 30 minutes at 37°C. Laemmli buffer was add to the mix and proteins were analyze by western blot.

Mass spectrometry analysis of FMRpolyG interactant and FMRpolyG N-terminus.

5x10exp6 HeLa cell were transfected with 18 µg of HA-FLAG tagged plasmid using Fugen HD (Promega) for 24h hours. Proteins were purified by HA-FLAG tandem purification kit according to the manufacturer's instruction (Sigma-Aldrich), separated on 4-12% bis-Tris Gel (NuPAGE) and visualized by silver staining (SilverQuest, Invitrogen). Gel bands were

excised and subjected to manual in-gel reduction in 10 mM DTT in 100 mM NH4HCO3 (Sigma Aldrich) for 1 h at 57°C, alkylated for 45 min in the dark with 55 mM iodoacetamide in 100 mM NH4HCO3 (Sigma Aldrich), washed in 25 mM NH4HCO3, dehydrated with acetonitrile and dried in SpeedVac 5301 Concentrator (Eppendorf). Then the gel pieces were rehydrated with 12.5 ng/µL trypsin or LysC solution (Promega) in 50 mM NH4HCO3 and incubated overnight at 37°C. The peptides were extracted twice with acetonitrile/water/formic acid-45/45/10-v/v/v followed by a final extraction with acetonitrile /formic acid (FA)-95/05-v/v. Extracted peptides were then analyzed using an Ultimate 3000 nano-RSLC (Thermo Scientific) coupled in line with an Orbitrap ELITE (Thermo Scientific). Peptides were separated on a C18 nano-column with a linear gradient of acetonitrile and analyzed with in a Top 20 CID (Collision-induced dissociation) data-dependent mass spectrometry with an inclusion list. Data were processed by database searching using SequestHT (Thermo Fisher Scientific) with Proteome Discoverer 1.4 software (Thermo Fisher Scientific) against a homemade database of all potential three frames translated proteins or peptides from the 5'UTR of *FMR1*. Precursor and fragment mass tolerance were set at 7 ppm and 0.5 Da respectively. Oxidation (M) and Nterminal Acetylation were set as variable modification, and Carbamidomethylation (C) as fixed modification. Peptides were filtered with the Fixed value node of Proteome Discoverer 1.4. Similarly, for identification of protein interactants, 5x10exp6 Neuro2A cell were transfected with 15 µg of FMRpolyG FLAG-HA double-tagged plasmid using Fugen HD (Promega) for 24h hours. Proteins were purified by HA-FLAG tandem purification kit according to the manufacturer's instruction (Sigma-Aldrich). Proteins were visualized by silver staining (SilverQuest, Invitrogen) after separation of 4-12% bis-Tris Gel (NuPAGE) and identified using NanoESI_Ion Trap (LTQ XL Thermo Fisher).

Immunofluorescence and immunohistochemistry.

Mouse or human brain sections were deparaffinized two times for 20 min in Histosol Plus (Shandon) and dehydrated as follows: twice in ethanol 100% (5 min), twice in ethanol 95% (5 min), once in ethanol 80% (5 min), once in ethanol 70% (5 min) and rinsed in PBS. Glass coverslips containing plated cells or brain sections treated as described above were fixed in PFA during 10 min and washed three times with PBS. The coverslips or slides were incubated for 10 min in PBS plus 0.5% Triton X-100 and washed three times with PBS before incubation during 1 hours with primary antibody against FMRpolyA (5FM, 1/100), FMRpolyG (8FM or 9FM, 1/50 to 1/100), ubiquitin (DAKO, Z0458), GFP (Abcam ab1218), Lap2 (Millipore 06-1002; Abcam ab185718; Abcam ab189993), Lamin B1 (Abcam

ab16048). Slides or coverslips were washed twice with PBS before incubation with a goat anti-rabbit or goat anti-mouse secondary antibody conjugated with Cyanine 3 (1/500 dilution; Fisher) for 60 min; incubated for 2 min in PBS 1X-DAPI (1/10 000 dilution) and rinsed twice with PBS 1X before mounting in Pro-Long media (Molecular Probes). Slides were examined using a fluorescence microscope (Leica). For immunochemistry, brain sections were deparaffinized followed by antigen retrieval using microwave treatment in 0.01 m sodium citrate and treatment with 10 µg/ml protein kinase for 20 min at 37°c. Endogenous peroxidase activity was blocked, and immunostaining was performed overnight at 4°C using antibody against Iba1 (Abcam ab15690), Gfap (Abcam ab7260), ubiquitin (Dako Z0458; 1:250), FMRpolyG (8FM or 9FM, 1:10 to 1/50) or Lap2 (Millipore 06-1002; Abcam ab185718; Abcam ab189993). Antigen–antibody complexes were visualized by incubation with DAB substrate (Dako) and slides were counterstained with hematoxylin and eosin.

RNA FISH coupled to immunofluorescence.

Mouse brain sections were deparaffinized and rehydrated. Coverslips containing primary culture of E18 mouse cortical neurons cells or brain sections were fixed in PFA during 10 min and washed three times with PBS. The coverslips or slides were incubated for 10 min in PBS plus 0.5% Triton X-100 and washed three times with PBS before pre-hybridization in 40% DMSO, 40% formamide, 10% BSA (10 mg/ml), 2 × SCC for 30 min. The coverslips or slides were hybridized for 2 h in 40% formamide, 10% DMSO, 2 × SCC, 2 mM vanadyl ribonucleoside, 60 µg/ml tRNA, 30 µg/ml BSA plus 0.75 µg (CCG)8x-Cy3 DNA oligonucleotide probe (Sigma). The coverslips or slides were washed twice in 2 × SCC/50% formamide and twice in 2 × SCC. The coverslips were incubated for 2 min in 2 × SCC/DAPI (1/10 000 dilution) and rinsed twice in 2 × SSC before mounting in Pro-Long media (Molecular Probes). Coverslips were examined using a fluorescence microscope (Leica). For FISH followed by immunofluorescence, after 2 × SCC wash, the slide were washed twice in PBS 1X. The slides were incubated 1 hour with primary antibody against FMRpolyGly antibody (8FM, 1/50). Slides were washed twice with PBS before incubation with a goat anti-mouse secondary antibody conjugated with cyanine-3 (1/500 dilution; Fisher) for 60 min; incubated for 2 min in PBS 1X-DAPI (1/10 000 dilution) and rinsed twice in PBS 1X before mounting in Pro-Long media (Molecular Probes). Slides were examined using a fluorescence microscope (Leica).

***Drosophila* models of FMRpolyG**

All Drosophila lines were maintained on standard culture and food conditions at 25^{o}C, while all crosses and experiments were performed at 29^{o}C. Control and driver lines used in this study are *w*^{1118} (control) from Bloomington, *Actin5C-GAL4/CyO* driver (ubiquitous driver line) as a gift from Zhe Han's lab, and RU486-inducible Geneswitch tubulin driver line *Tub5-GAL4* (ubiquitous expression) as a gift from Scott Pletcher's lab. DNA fragments containing FMRpolyG-GFP or polyG-GFP without the C-terminal sequence were PCR amplified from counterpart of mammalian transfection vectors described elsewhere and inserted to a pUAST vector between EcoRI and XbaI sites. All constructs were sequence verified by Sanger sequencing and transgenic flies with these constructs were made via standard p-element insertion (BestGene, CA). Transgene expression levels of GFP gene were analyzed 3 days after induction with RU486 in flies from the individual lines crossed to the *Tub5-GAL4* driver, and those with comparable RNA expression levels were used for this study. The fly eclosion assay has been described elsewhere (Todd et al, 2013). Briefly, homozygous UAS transgenic lines or control lines were crossed to a *Actin5C-GAL4* ubiquitous driver line balanced over a marker chromosome (CyO), on standard food at 29^{o}C, if the transgene elicited no toxicity, then the number of progeny bearing the GAL4 driver would be expected to be equivalent to those bearing the CyO marker. Over 100 flies of each genotype were scored from multiple crosses. The ratio of expected progeny carrying the transgene compared to those carrying the CyO marker was expressed as a percentage of the expected ratio of one. These percentages were then compared using a Fischer exact test to determine statistical significance. For the fly survival assay, The UAS transgenic lines or control lines were crossed to *Tub5-Gal4* geneswitch driver flies on standard food absent of RU486 at 29^{o}C. Adult offspring of the desired genotypes were collected 2-3 days after eclosion and transferred to standard fly food containing 200 μM RU486 without yeast granules. The flies were transferred to fresh food with drug every 2-3 days. Each genotype started with at least 4 vials of 25 flies/vial and the survival was determined daily or every other day for 3 weeks.

Mouse models, genotyping and phenotyping.

All mouse procedures were done according protocols approved by the Committee on Animal Resources of the ICS animal facility and under the French and European authority guidelines.

For transgene construction, human 5'UTR FMR1 fused in the glycine frame to the GFP (Addgene) was cloned between the FseI and SmaI sites of the Rosa26 5' arm – CAG promoter – LOXP - SV40 polyA 3x - LOXP – Rosa26 3' arm vector Ai2 (Addgene). Deleted transgene for the ACG near-cognate codon was constructed by deleting 100 nucleotides from

FseI to KasI sites within the *FMR1* 5'UTR plasmid. Both mutant mouse lines were established at the ICS (Mouse Clinical Institute; http://www-mci.u-strasbg.fr). Both linearized constructs were electroporated separately in C57Bl/l6N mouse embryonic stem (ES) cells. After G418 selection, targeted clones were identified by long-range PCR using external primers and further confirmed by Southern blot with an internal probe against Neomycin and 5' external probe against Rosa26. Two positive ES clones for each future transgenic mice were injected into Balb/cN blastocysts. Resulting male chimeras were bred with wild type C57Bl/6N females to obtain germline transmission. Deletion of the floxed STOP cassette were performed by breeding F1 males with CMV-cre deleter females (Birling et al., 2012) or Nestin-cre delete mice (Isaka et al., 1999). Genotyping across the expanded CGG repeats was performed using the Expand High Fidelity PCR System (Roche, 11-732 -650 001) according manufacturer instructions and supplemented with 2,5 M Betaine (B0300 Sigma, 12,5 µl of 5 M Betaine for a final PCR volume of 25 µl) with one denaturation step at 94 °C for 2 min, 30 cycles of amplification 94 °C for 1 min, 60 °C for 1 min, 72 °C for 2 min and a final step at 72 °C for 5 min using the forward primer 5'-TCGACCTGCAGCCCAAGCTAGATCG and the reverse primer 5'-TCCTTGAAGAAGATGGTGCGCTCC. Rotarod test (Bioseb, Chaville, France) was performed with three testing trials during which the rotation speed accelerated from 4 to 40 rpm in 5 min. Trials were separated by 5-10 min interval. The average latency was used as index of motor coordination performance. Grip test: this test measures the maximal muscle strength (g) using an isometric dynamometer connected to a grid (Bioseb). Mice were allowed to grip the grid with all its paws then they were pulled backwards until they released it. Each mouse was submitted to 3 consecutive trials immediately after the modified SHIRPA procedure. The maximal strength developed by the mouse before releasing the grid was recorded and the average value of the three trials adjusted to body weight. The string test consisted of 3 trials separated by 5-10 min interval. On each trial the forepaws of the animal were placed on the thread that is a wire stretched horizontally 40 cm above the bench. The latency the animal took to catch the wire with its hindpaws was recorded. Open field test: mice were tested in automated open fields (Panlab, Barcelona, Spain), each virtually divided into central and peripheral regions. The open fields were placed in a room homogeneously illuminated at 150 Lux. Each mouse was placed in the periphery of the open field and allowed to explore freely the apparatus for 30 min, with the experimenter out of the animal's sight. The distance traveled, the number of rears, and time spent in the central and peripheral regions were recorded over the test session. The number of entries and the percent time spent in center area are used as index of emotionality/anxiety

Quantitative real time RT-PCR.

Total RNAs from mouse tissues or cells were isolated by TriReagent (Molecular Research Center). cDNAs were generated using the Transcriptor High Fidelity cDNA synthesis kit (Roche Diagnostics) for quantification of mRNAs. qPCR of mRNAs were realized using the LightCycler 480 SYBR Green I Master (Roche) in a Lightcycler 480 (Roche) with 15 min at 94°C followed by 50 cycles of 15 sec at 94°C, 20 sec at 58°C and 20 sec at 72°C. RPLPO mRNA was used as standard and data were analyzed using the Lightcycler 480 analysis software (2ΔCt method).

Subcellular fractionation and PCR

Cells were scraped in PBS, cells were pelleted by centrifugation at 3000 rpm for 10 min minutes. The pellet was resuspended in Dautry Buffer (Tris HCl pH 7,8, 10 mM; NaCl 140 mM, MgCl2 1,5 mM; EDTA 10 mM, NP40 0,5%) and kept on ice 5 minutes. The homogenate was centrifuged at 3000 rpm for 15 minutes during 5 min at 4°C, pellet corresponding to nuclear fraction and supernatant to cytosolic fraction. Cytosolic fraction was centrifuged at 13000 rpm for 15 minutes to remove potential nucleus and 1 ml TriReagent (Molecular Research Center) was added. The pellet was washed with 400 µl of Dautry Buffer, centrifuged at 3000 rpm for 5 minutes at 4°C. Supernatant was removed and the pellet was homogenized in 400 µl of Dautry Buffer and 1 ml TriReagent (Molecular Research Center) was added. Total RNA from nuclear or cytosolic fraction was isolated as described in the manufacturer's protocol of TriReagent (Molecular Research Center). cDNAs were generated using the Transcriptor High Fidelity cDNA synthesis kit (Roche Diagnostics) for quantification of mRNAs. PCR was performed with Taq polymerase (Roche), one denaturation step at 94 °C for 2 min, 25 cycles of amplification 94 °C for 1 min, 60 °C for 1 min, 72 °C for 2 min and a final step at 72 °C for 5 min using the primer described below. The PCR products were precipitated, analyzed by electrophoresis on a 6.5% polyacrylamide gel, stained with ethidium bromide and quantified with a Typhoon scanner.

Oligonucleotides.

RPLO_FW	GAAGTCACTGTGCCAGCCCA
RPLO_REV	GAAGGTGTAATCCGTCTCCA
U6_FW	CTCGCTTCGGCAGCACATATA
U6_REV	GGAACGCTTCACGAATTTGCG

FMR1_FW	GAAAACAACTGGCAGCCTGA
FMR1_REV	AGCTAACCACCAACAGCAAG
GFP_FW	ACGTAAACGGCCACAAGTTC
GFP_REV	AAGTCGTGCTGCTTCATGTG
(CGG)60x_FW	GAACCCACTGCTTACTGGCTTA
(CGG)60x_REV	AACGCTAGCCAGTTGGGTC
Transgene mouse FMRpolyG_FW	GCAAGCTGACCCTGAAGTTC
Transgene mouse FMRpolyG_REV	GTCTTGTAGTTGCCGTCGTC

ACKNOWLEDGMENTS.

We thank Prof. PJ Hagerman (Dpt. of Biochemistry and Molecular Medicine, School of Medicine, Davis, USA) for the kind gift of human samples. We thank all IGBMC and ICS common core facilities for assistance.

FUNDING ACKNOWLEDGMENTS.

This work was supported by ANR-12-RARE-0001 E-RARE "CURE FXTAS" (NCB, RKH) and ANR-14-RARE-0003 E-RARE "Drug_FXSPreMut" (NCB, RW, CM), ERC-2012-StG #310659 "RNA DISEASES"; ANR-14-CE10-0016-01 "MITO-FXTAS" (NCB); ANR-10-INBS-07 PHENOMIN (ICS), ANR-10-LABX-0030-INRT (IGBMC) and ANR-10-IDEX-0002-02 (IGBMC); Hersenstichting Nederland F2012(1)-101 (RW); Hersenstichting Nederland F2015(1)-02 (RKH); National Institutes of Health NINDS RO1 NS079775 (RW); NICHD HD040661 (PJH), NIH R01NS08681001 (PKT) and the Department of Veterans Affairs BLRD 1I21BX001841 (PKT).

AUTHOR CONTRIBUTIONS.

Experiments were performed by C.S., RA.B., F.H., PS., L.J., P.T., A.G., H.J., H.M., A.V., MF.C., M.WD., MC.B. M.O.A., M.J. and F.R. Clinical samples and patient data were obtained from M.A., R.W., V.M.C, F.T. and PJ.H. Data were collected and analysed by C.S., F.H., T.S., G.P., M.J., R.K.H., R.W., S.V., C.M., PK.T. and N.C.B. The study was designed, coordinated and written by C.S. and N.C.B with editorial inputs by all authors.

CONFLICT OF INTEREST.

PKT serves as a consultant for Denali Therapeutics. This consultation role does not relate to the work described here

REFERENCES:

Apartis E, Blancher A, Meissner WG, Guyant-Maréchal L, Maltête D, De Broucker T, Legrand AP, Bouzenada H, Thanh HT, Sallansonnet-Froment M, Wang A, Tison F, Roué-Jagot C, Sedel F, Charles P, Whalen S, Héron D, Thobois S, Poisson A, Lesca G, Ouvrard-Hernandez AM, Fraix V, Palfi S, Habert MO, Gaymard B, Dussaule JC, Pollak P, Vidailhet M, Durr A, Barbot JC, Gourlet V, Brice A, Anheim M. FXTAS: new insights and the need for revised diagnostic criteria. Neurology. 2012 Oct 30;79(18):1898-907.

Arocena DG, Iwahashi CK, Won N, Beilina A, Ludwig AL, Tassone F, Schwartz PH, Hagerman PJ. (2005). Induction of inclusion formation and disruption of lamin A/C structure by premutation CGG-repeat RNA in human cultured neural cells. Hum Mol Genet. 14: 3661-71.

Ash PE1, Bieniek KF, Gendron TF, Caulfield T, Lin WL, Dejesus-Hernandez M, van Blitterswijk MM, Jansen-West K, Paul JW 3rd, Rademakers R, Boylan KB, Dickson DW, Petrucelli L. Unconventional translation of C9ORF72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. Neuron 2013;77: 639-646.

Aspden JL, Eyre-Walker YC, Phillips RJ, Amin U, Mumtaz MA, Brocard M, Couso JP. Extensive translation of small Open Reading Frames revealed by Poly-Ribo-Seq. Elife. 2014 Aug 21;3:e03528.

Barbosa C, Peixeiro I, Romão L. Gene expression regulation by upstream open reading frames and human disease. PLoS Genet. 2013;9(8):e1003529.

Birling MC, Dierich A, Jacquot S, Hérault Y, Pavlovic G. Highly-efficient, fluorescent, locus directed cre and FlpO deleter mice on a pure C57BL/6N genetic background. Genesis. 2012 Jun;50(6):482-9.

Boissart C, Poulet A, Georges P, Darville H, Julita E, Delorme R, Bourgeron T, Peschanski M, Benchoua A. Differentiation from human pluripotent stem cells of cortical neurons of the superficial layers amenable to psychiatric disease modeling and high-throughput drug screening. Transl Psychiatry. 2013 Aug 20;3:e294.

Buijsen RA, Visser JA, Kramer P, Severijnen EA, Gearing M, Charlet-Berguerand N, Sherman SL, Berman RF, Willemsen R, Hukema RK. Presence of inclusions positive for polyglycine containing protein, FMRpolyG, indicates that repeat-associated non-AUG translation plays a role in fragile X-associated primary ovarian insufficiency. Hum Reprod. 2016 Jan;31(1):158-68.

Brouwer JR, Huizer K, Severijnen LA, Hukema RK, Berman RF, Oostra BA, Willemsen R (2008) CGG repeat length and neuropathological and molecular correlates in a mouse model for fragile X-associated tremor/ataxia syndrome. *J Neurochem* **107**: 1671-1682

Cai M1, Huang Y, Ghirlando R, Wilson KL, Craigie R, Clore GM. Solution structure of the constant region of nuclear envelope protein LAP2 reveals two LEM-domain structures: one binds BAF and the other binds DNA. EMBO J. 2001 Aug 15;20(16):4399-407.

Calvo SE, Pagliarini DJ, Mootha VK. Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. Proc Natl Acad Sci U S A. 2009 May 5;106(18):7507-12.

Cleary JD, Ranum LP. Repeat associated non-ATG (RAN) translation: new starts in microsatellite expansion disorders. Curr Opin Genet Dev. 2014 Jun;26:6-15.

DeJesus-Hernandez M, Mackenzie IR, Boeve BF, Boxer AL, Baker M, Rutherford NJ, Nicholson AM, Finch NA, Flynn H, Adamson J, Kouri N, Wojtas A, Sengdy P, Hsiung GY, Karydas A, Seeley WW, Josephs KA, Coppola G, Geschwind DH, Wszolek ZK, Feldman H, Knopman DS, Petersen RC, Miller BL, Dickson DW, Boylan KB, Graff-Radford NR, Rademakers R. (2011). Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS. Neuron. 72:245-56.

Dorner D, Vlcek S, Foeger N, Gajewski A, Makolm C, Gotzmann J, Hutchison CJ, Foisner R. Lamina-associated polypeptide 2alpha regulates cell cycle progression and differentiation via the retinoblastoma-E2F pathway. J Cell Biol. 2006 Apr 10;173(1):83-93.

Dubińska-Magiera M1, Chmielewska M, Kozioł K, Machowska M, Hutchison CJ, Goldberg MW, Rzepecki R. Xenopus LAP2β protein knockdown affects location of lamin B and nucleoporins and has effect on assembly of cell nucleus and cell viability. Protoplasma. 2015 Jul 25. [Epub ahead of print]

Entezam A, Biacsi R, Orrison B, Saha T, Hoffman GE, Grabczyk E, Nussbaum RL, Usdin K. (2007). Regional FMRP deficits and large repeat expansions into the full mutation range in a new Fragile X premutation mouse model. Gene. 395:125-34.

Feng Y, Zhang F, Lokey LK, Chastain JL, Lakkis L, Eberhart D, Warren ST. Translational suppression by trinucleotide repeat expansion at FMR1. Science. 1995 May 5;268(5211):731-4.

Freibaum BD, Lu Y, Lopez-Gonzalez R, Kim NC, Almeida S, Lee KH, Badders N, Valentine M, Miller BL, Wong PC, Petrucelli L, Kim HJ, Gao FB, Taylor JP. GGGGCC repeat expansion in C9orf72 compromises nucleocytoplasmic transport. Nature. 2015 Sep 3;525(7567):129-33.

Fritsch C, Herrmann A, Nothnagel M, Szafranski K, Huse K, Schumann F, Schreiber S, Platzer M, Krawczak M, Hampe J, Brosch M. Genome-wide search for novel human uORFs and N-terminal protein extensions using ribosomal footprinting. Genome Res. 2012 Nov;22(11):2208-18.

Foisner R1, Gerace L. Integral membrane proteins of the nuclear envelope interact with lamins and chromosomes, and binding is modulated by mitotic phosphorylation. Cell. 1993 Jul 2;73(7):1267-79.

Furukawa K, Panté N, Aebi U, Gerace L. Cloning of a cDNA for lamina-associated polypeptide 2 (LAP2) and identification of regions that specify targeting to the nuclear envelope. EMBO J. 1995 Apr 18;14(8):1626-36.

Gant TM, Harris CA, Wilson KL. Roles of LAP2 proteins in nuclear assembly and DNA replication: truncated LAP2beta proteins alter lamina assembly, envelope formation, nuclear

size, and DNA replication efficiency in Xenopus laevis extracts. J Cell Biol. 1999 Mar 22;144(6):1083-96.

Greco CM, Hagerman RJ, Tassone F, Chudley AE, Del Bigio MR, Jacquemont S, Leehey M, Hagerman PJ (2002). Neuronal intranuclear inclusions in a new cerebellar tremor/ataxia syndrome among fragile X carriers. Brain 125:1760-1771.

Greco CM, Berman RF, Martin RM, Tassone F, Schwartz PH, Chang A, Trapp BD, Iwahashi C, Brunberg J, Grigsby J, Hessl D, Becker EJ, Papazian J, Leehey MA, Hagerman RJ, Hagerman PJ. (2006) Neuropathology of fragile X-associated tremor/ataxia syndrome (FXTAS). *Brain*. **1**:243-55.

Hagerman RJ, Leehey M, Heinrichs W, Tassone F, Wilson R, Hills J, Grigsby J, Gage B, Hagerman PJ. (2001). Intention tremor, parkinsonism, and generalized brain atrophy in male carriers of fragile X. Neurology. 1:127-30.

Hagerman and Hagerman (2004) The fragile-X premutation: a maturing perspective. *Am J Hum Genet* **74**: 805-816

Hashem V, Galloway JN, Mori M, Willemsen R, Oostra BA, Paylor R, Nelson DL. (2009). Ectopic expression of CGG containing mRNA is neurotoxic in mammals. Hum Mol Genet. 18: 2443-51.

Hoem G, Raske CR, Garcia-Arocena D, Tassone F, Sanchez E, Ludwig AL, Iwahashi CK, Kumar M, Yang JE, Hagerman PJ. (2011) CGG-repeat length threshold for FMR1 RNA pathogenesis in a cellular model for FXTAS. Hum Mol Genet. 20:2161-70.

Hukema RK, Buijsen RA, Raske C, Severijnen LA, Nieuwenhuizen-Bakker I, Minneboo M, Maas A, de Crom R, Kros JM, Hagerman PJ, Berman RF, Willemsen R. Induced expression of expanded CGG RNA causes mitochondrial dysfunction in vivo. Cell Cycle. 2014;13(16):2600-8.

Hukema RK, Buijsen RA, Schonewille M, Raske C, Severijnen LA, Nieuwenhuizen-Bakker I, Verhagen RF, van Dessel L, Maas A, Charlet-Berguerand N, De Zeeuw CI, Hagerman PJ,

Berman RF, Willemsen R. Reversibility of neuropathology and motor deficits in an inducible mouse model for FXTAS. Hum Mol Genet. 2015 Sep 1;24(17):4948-57.

Ingolia NT, Lareau LF, Weissman JS. Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. Cell. 2011 Nov 11;147(4):789-802.

Isaka F, Ishibashi M, Taki W, Hashimoto N, Nakanishi S, Kageyama R. Ectopic expression of the bHLH gene Math1 disturbs neural development. Eur J Neurosci. 1999; 11:2582-8;

Iwahashi CK, Yasui DH, An HJ, Greco CM, Tassone F, Nannen K, Babineau B, Lebrilla CB, Hagerman RJ, Hagerman PJ. (2006). Protein composition of the intranuclear inclusions of FXTAS. Brain 129: 256-271

Jin P, Zarnescu DC, Zhang F, Pearson CE, Lucchesi JC, Moses K, Warren ST. (2003). RNA-mediated neurodegeneration caused by the fragile X premutation rCGG repeats in Drosophila. Neuron 39:739-747.

Jin P, Duan R, Qurashi A, Qin Y, Tian D, Rosser TC, Liu H, Feng Y, Warren ST. (2007). Pur alpha binds to rCGG repeats and modulates repeat-mediated neurodegeneration in a Drosophila model of fragile X tremor/ataxia syndrome. Neuron 55: 556-564

Jacquemont S, Hagerman RJ, Leehey M, Grigsby J, Zhang L, Brunberg JA, Greco C, Des Portes V, Jardini T, Levine R, Berry-Kravis E, Brown WT, Schaeffer S, Kissel J, Tassone F, Hagerman PJ. (2003). Fragile X premutation tremor/ataxia syndrome: molecular, clinical, and neuroimaging correlates. Am J Hum Genet. 4:869-78.

Jacquemont S, Hagerman RJ, Leehey MA, Hall DA, Levine RA, Brunberg JA, Zhang L, Jardini T, Gane LW, Harris SW, Herman K, Grigsby J, Greco CM, Berry-Kravis E, Tassone F, Hagerman PJ. (2004) Penetrance of the fragile X-associated tremor/ataxia syndrome in a premutation carrier population. *JAMA*. **4**:460-9.

Jacquemont S, Leehey MA, Hagerman RJ, Beckett LA, Hagerman PJ. Size bias of fragile X premutation alleles in late-onset movement disorders. J Med Genet. 2006 Oct;43(10):804-9.

Ji Z, Song R, Regev A, Struhl K. Many lncRNAs, 5'UTRs, and pseudogenes are translated and some are likely to express functional proteins. Elife. 2015 Dec 19;4. pii: e08890.

Jovičić A, Mertens J, Boeynaems S, Bogaert E, Chai N, Yamada SB, Paul JW 3rd, Sun S, Herdy JR, Bieri G, Kramer NJ, Gage FH, Van Den Bosch L, Robberecht W, Gitler AD. Modifiers of C9orf72 dipeptide repeat toxicity connect nucleocytoplasmic transport defects to FTD/ALS. Nat Neurosci. 2015 Sep;18(9):1226-9.

Jung L, Tropel P, Moal Y, Teletin M, Jeandidier E, Gayon R, Himmelspach C, Bello F, André C, Tosch A, Mansouri A, Bruant-Rodier C, Bouillé P, Viville S. ONSL and OSKM cocktails act synergistically in reprogramming human somatic cells into induced pluripotent stem cells. Mol Hum Reprod. 2014 Jun;20(6):538-49.

Leehey MA, Berry-Kravis E, Goetz CG, Zhang L, Hall DA, Li L, Rice CD, Lara R, Cogswell J, Reynolds A, Gane L, Jacquemont S, Tassone F, Grigsby J, Hagerman RJ, Hagerman PJ. (2008) FMR1 CGG repeat length predicts motor dysfunction in premutation carriers. *Neurology*. **16**:1397-402.

Liu, J., Koscielska, K. A., Cao, Z., Hulsizer, S., Grace, N., Mitchell, G., Nacey, C., Githinji, J., Mcgee, J., Garcia-Arocena, D., Hagerman, R. J., Nolta, J., Pessah, I. N., Hagerman, P. J. (2012). Signaling defects in iPSC-derived fragile X premutation neurons. Hum Mol Genet 21(17):3795-805.

Mackenzie IR, Frick P, Grässer FA, Gendron TF, Petrucelli L, Cashman NR, Edbauer D, Kremmer E, Prudlo J, Troost D, Neumann M. Quantitative analysis and clinico-pathological correlations of different dipeptide repeat protein pathologies in C9ORF72 mutation carriers. Acta Neuropathol. 2015 Dec;130(6):845-61.

Marteyn A, Maury Y, Gauthier MM, Lecuyer C, Vernet R, Denis JA, Pietu G, Peschanski M, Martinat C. Mutant human embryonic stem cells reveal neurite and synapse formation defects in type 1 myotonic dystrophy. Cell Stem Cell. 2011 Apr 8;8(4):434-44.

Mori K, Weng SM, Arzberger T, May S, Rentzsch K, Kremmer E, Schmid B, Kretzschmar HA, Cruts M, Van Broeckhoven C, Haass C, Edbauer D. The C9orf72 GGGGCC repeat is

translated into aggregating dipeptide-repeat proteins in FTLD/ALS. Science. 2013 Mar 15;339(6125):1335-8.

Naetar N, Korbei B, Kozlov S, Kerenyi MA, Dorner D, Kral R, Gotic I, Fuchs P, Cohen TV, Bittner R, Stewart CL, Foisner R. Loss of nucleoplasmic LAP2alpha-lamin A complexes causes erythroid and epidermal progenitor hyperproliferation. Nat Cell Biol. 2008 Nov;10(11):1341-8.

Nili E, Cojocaru GS, Kalma Y, Ginsberg D, Copeland NG, Gilbert DJ, Jenkins NA, Berger R, Shaklai S, Amariglio N, Brok-Simoni F, Simon AJ, Rechavi G. Nuclear membrane protein LAP2beta mediates transcriptional repression alone and together with its binding partner GCL (germ-cell-less). J Cell Sci. 2001 Sep;114(Pt 18):3297-307.

Oh SY, He F, Krans A, Frazer M, Taylor JP, Paulson HL, Todd PK. RAN translation at CGG repeats induces ubiquitin proteasome system impairment in models of fragile X-associated tremor ataxia syndrome. Hum Mol Genet. 2015 Aug 1;24(15):4317-26.

Primerano B, Tassone F, Hagerman RJ, Hagerman P, Amaldi F, Bagni C. (2002) Reduced FMR1 mRNA translation efficiency in fragile X patients with premutations. *RNA*. **12**:1482-8.

Qin M, Entezam A, Usdin K, Huang T, Liu ZH, Hoffman GE, Smith CB. A mouse model of the fragile X premutation: effects on behavior, dendrite morphology, and regional rates of cerebral protein synthesis. Neurobiol Dis. 2011 Apr;42(1):85-98.

Renton AE, Majounie E, Waite A, Simón-Sánchez J, Rollinson S, Gibbs JR, Schymick JC, Laaksovirta H, van Swieten JC, Myllykangas L, Kalimo H, Paetau A, Abramzon Y, Remes AM, Kaganovich A, Scholz SW, Duckworth J, Ding J, Harmer DW, Hernandez DG, Johnson JO, Mok K, Ryten M, Trabzuni D, Guerreiro RJ, Orrell RW, Neal J, Murray A, Pearson J, Jansen IE, Sondervan D, Seelaar H, Blake D, Young K, Halliwell N, Callister JB, Toulson G, Richardson A, Gerhard A, Snowden J, Mann D, Neary D, Nalls MA, Peuralinna T, Jansson L, Isoviita VM, Kaivorinne AL, Hölttä-Vuori M, Ikonen E, Sulkava R, Benatar M, Wuu J, Chiò A, Restagno G, Borghero G, Sabatelli M; ITALSGEN Consortium, Heckerman D, Rogaeva E, Zinman L, Rothstein JD, Sendtner M, Drepper C, Eichler EE, Alkan C, Abdullaev Z, Pack

SD, Dutra A, Pak E, Hardy J, Singleton A, Williams NM, Heutink P, Pickering-Brown S, Morris HR, Tienari PJ, Traynor BJ. A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. Neuron. 2011 Oct 20;72(2):257-68.

Sellier C, Rau F, Liu Y, Tassone F, Hukema RK, Gattoni R, Schneider A, Richard S, Willemsen R, Elliott DJ, Hagerman PJ, Charlet-Berguerand N. (2010). Sam68 sequestration and partial loss of function are associated with splicing alterations in FXTAS patients. EMBO J. 29: 1248-61.

Sellier, C., Freyermuth, F., Tabet, R., Tran, T., He, F., Ruffenach, F., Alunni, V., Moine, H., Thibault, C., Page, A., Tassone, F., Willemsen, R., Disney, M. D., Hagerman, P. J., Todd, P. K., Charlet-Berguerand, N. (2013). Sequestration of DROSHA and DGCR8 by expanded CGG RNA repeats alters microRNA processing in fragile X-associated tremor/ataxia syndrome. Cell Rep 3(3):869-80.

Seltzer MM1, Baker MW, Hong J, Maenner M, Greenberg J, Mandel D. Prevalence of CGG expansions of the FMR1 gene in a US population-based sample. Am J Med Genet B Neuropsychiatr Genet. 2012 Jul;159B(5):589-97.

Shumaker DK1, Lee KK, Tanhehco YC, Craigie R, Wilson KL. LAP2 binds to BAF.DNA complexes: requirement for the LEM domain and modulation by variable regions. EMBO J. 2001 Apr 2;20(7):1754-64.

Somech R, Shaklai S, Geller O, Amariglio N, Simon AJ, Rechavi G, Gal-Yam EN. The nuclear-envelope protein and transcriptional repressor LAP2beta interacts with HDAC3 at the nuclear periphery, and induces histone H4 deacetylation. J Cell Sci. 2005 Sep 1;118(Pt 17):4017-25.

Sofola OA, Jin P, Qin Y, Duan R, Liu H, de Haro M, Nelson DL, Botas J. (2007). RNA-binding proteins hnRNP A2/B1 and CUGBP1 suppress fragile X CGG premutation repeat-induced neurodegeneration in a Drosophila model of FXTAS. Neuron 55: 565-571.

Sonenberg N, Hinnebusch AG. Regulation of translation initiation in eukaryotes: mechanisms and biological targets. Cell. 2009 Feb 20;136(4):731-45.

Tassone F, Hagerman RJ, Loesch DZ, Lachiewicz A, Taylor AK, Hagerman PJ. (2000). Fragile X males with unmethylated, full mutation trinucleotide repeat expansions have elevated levels of FMR1 messenger RNA. Am J Med Genet 94: 232–236.

Tassone F, Iwahashi C, Hagerman PJ. (2004). FMR1 RNA within the intranuclear inclusions of fragile X-associated tremor/ataxia syndrome (FXTAS). RNA Biol 1:103-5.

Tassone F, Adams J, Berry-Kravis EM, Cohen SS, Brusco A, Leehey MA, Li L, Hagerman RJ, Hagerman PJ. (2007). CGG repeat length correlates with age of onset of motor signs of the fragile X-associated tremor/ataxia syndrome (FXTAS). Am J Med Genet B Neur Genet. 4:566-9.

Tassone F, Beilina A, Carosi C, Albertosi S, Bagni C, Li L, Glover K, Bentley D, Hagerman PJ. Elevated FMR1 mRNA in premutation carriers is due to increased transcription. RNA. 2007;13(4):555-62

Tassone F, Iong KP, Tong TH, Lo J, Gane LW, Berry-Kravis E, Nguyen D, Mu LY, Laffin J, Bailey DB, Hagerman RJ. FMR1 CGG allele size and prevalence ascertained through newborn screening in the United States. Genome Med. 2012 Dec 21;4(12):100.

Todd, P. K., Oh, S. Y., Krans, A., He, F., Sellier, C., Frazer, M., Renoux, A. J., Chen, K. C., Scaglione, K. M., Basrur, V., Elenitoba-Johnson, K., Vonsattel, J. P., Louis, E. D., Sutton, M. A., Taylor, J. P., Mills, R. E., Charlet-Berguerand, N., Paulson, H. L. (2013). CGG repeat-associated translation mediates neurodegeneration in fragile X tremor ataxia syndrome. Neuron 78(3):440-55.

Willemsen R, Hoogeveen-Westerveld M, Reis S, Holstege J, Severijnen LA, Nieuwenhuizen IM, Schrier M, van Unen L, Tassone F, Hoogeveen AT, Hagerman PJ, Mientjes EJ, Oostra BA. (2003). The FMR1 CGG repeat mouse displays ubiquitin-positive intranuclear neuronal inclusions; implications for the cerebellar tremor/ataxia syndrome. Hum Mol Genet 12: 949-959.

Yang, W. Y., Wilson, H. D., Velagapudi, S. P., Disney, M. D. (2015). Inhibition of Non-ATG Translational Events in Cells via Covalent Small Molecules Targeting RNA. J Am Chem Soc 137(16):5336-45.

Zhang K, Donnelly CJ, Haeusler AR, Grima JC, Machamer JB, Steinwald P, Daley EL, Miller SJ, Cunningham KM, Vidensky S, Gupta S, Thomas MA, Hong I, Chiu SL, Huganir RL, Ostrow LW, Matunis MJ, Wang J, Sattler R, Lloyd TE, Rothstein JD. The C9orf72 repeat expansion disrupts nucleocytoplasmic transport. Nature. 2015 Sep 3;525(7567):56-61

Zhang YJ, Gendron TF, Grima JC, Sasaguri H, Jansen-West K, Xu YF, Katzman RB, Gass J, Murray ME, Shinohara M, Lin WL, Garrett A, Stankowski JN, Daughrity L, Tong J, Perkerson EA, Yue M, Chew J, Castanedes-Casey M, Kurti A, Wang ZS, Liesinger AM, Baker JD, Jiang J, Lagier-Tourenne C, Edbauer D, Cleveland DW, Rademakers R, Boylan KB, Bu G, Link CD, Dickey CA, Rothstein JD, Dickson DW, Fryer JD, Petrucelli L. C9ORF72 poly(GA) aggregates sequester and impair HR23 and nucleocytoplasmic transport proteins. Nat Neurosci. 2016 Mar 21. doi: 10.1038/nn.4272. [Epub ahead of print]

Zu, T., Gibbens, B., Doty, N. S., Gomes-Pereira, M., Huguet, A., Stone, M. D., Margolis, J., Peterson, M., Markowski, T. W., Ingram, M. A., Nan, Z., Forster, C., Low, W. C., Schoser, B., Somia, N. V., Clark, H. B., Schmechel, S., Bitterman, P. B., Gourdon, G., Swanson, M. S., Moseley, M., Ranum, L. P. (2011). Non-ATG-initiated translation directed by microsatellite expansions. Proc Natl Acad Sci U S A 108(1):260-5.

Zu, T., Liu, Y., Banez-Coronel, M., Reid, T., Pletnikova, O., Lewis, J., Miller, T. M., Harms, M. B., Falchook, A. E., Subramony, S. H., Ostrow, L. W., Rothstein, J. D., Troncoso, J. C., Ranum, L. P. (2013). RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. Proc Natl Acad Sci U S A 110(51):E4968-77.

Zullo JM, Demarco IA, Piqué-Regi R, Gaffney DJ, Epstein CB, Spooner CJ, Luperchio TR, Bernstein BE, Pritchard JK, Reddy KL, Singh H. DNA sequence-dependent compartmentalization and silencing of chromatin at the nuclear lamina. Cell. 2012 Jun 22;149(7):1474-87.

FIGURE LEGENDS:

Figure 1. Translation of CGG repeats initiates at an upstream near-cognate codon.
(**A**) Immunoblotting against GFP of HeLa cells transfected with expanded CGG repeats embedded or not in the 5'UTR of *FMR1* and fused in all three possible frames with the GFP deleted of its ATG. (**B**) Upper panel, schemes of the *FMR1* 5'UTR deletion constructs tested. Middle panel, immunoblotting against GFP of HeLa cells transfected with mutants of *FMR1* 5'UTR containing expanded CGG repeats fused to the GFP in the glycine frame. Lower panel, quantification of FMRpolyG-GFP expression reported to GAPDH. Error bars indicate s.e.m of 3 independent transfections. Student t-test, *** indicates $p<0.001$. (**C**) LC-MS/MS spectra of the N-terminal part of the immunoprecipitated and trypsin digested protein translated from expanded CGG embedded in the 5'UTR of *FMR1*. (**D**) Nucleotide sequence of part of the human 5'UTR spliced to exon1 of *FMR1* with 29 CGG repeats. Amino acid sequence of *FMR1* uORF translated into FMRpolyG is indicated in red with a red arrow highlighting the ACG translation start. Amino acid sequence of the beginning of the FMRP ORF is indicated in green.

Figure 2. A minimum of 60 expanded CGG repeats is required to detect FMRpolyG.
(**A**) Upper panel, immunoblotting against the FLAG tag of HeLa cells transfected with various length of expanded CGG repeats embedded in the 5'UTR of *FMR1* and fused in the glycine frame with the FLAG tag. Middle panel, control immunoblotting against GAPDH. Lower panel, quantification of FMRpolyG-FLAG levels reported to GAPDH. (**B**) Upper panel, immunoblotting against GFP of HeLa cells transfected with various length of expanded CGG repeats embedded in the 5'UTR of *FMR1* and fused in the glycine frame with GFP. Middle panel, control immunoblotting against GAPDH. Lower panel, quantification of FMRpolyG-GFP levels reported to GAPDH. (**C**) Left panel, schemes of the 5'UTR of *FMR1* constructs with ACG mutations. FMRpolyG uORF is FLAG-tagged in the glycine frame, while FMRP ORF is GFP-tagged. Right upper and medium panels, immunoblotting against FLAG, GFP and GAPDH of HeLa cells transfected with mutants of the 5'UTR of *FMR1* containing expanded CGG repeats fused to the FLAG tag in the glycine frame, while the downstream FMRP ORF is fused to the GFP. Right lower panel, quantification of FMRpolyG-FLAG and FMRP-GFP levels reported to GAPDH. (**D**)

Immunofluorescence against the N-terminal part of FMRpolyG (red, 8FM antibody) and ubiquitin (green) on brain sections (hippocampal area) of FXTAS or control individuals. Scale bars, 10 μm. Nuclei were counterstained with DAPI. (**E**) Immunoblotting against FMRpolyG N-terminus (8FM antibody) of insoluble fraction of brain lysate of FXTAS and age-matched individuals. Number of expanded CGG in *FMR1* are indicated as # CGG. Error bars indicate s.e.m of 3 independent transfections. Student t-test, * indicates $p<0.05$, *** indicates $p<0.001$.

Figure 3. Expression of FMRpolyG is pathogenic to mice.

(**A**) Schemes of the mouse transgene constructs. The human full-length or deleted 5'UTR of *FMR1* are represented as red boxes, GFP ORF as green boxes, loxP sites as blue boxes, while promoters, polyadenylation sequences and start and stop of translation are indicated by arrow or black lines. (**B**) Quantitative RT-PCR analysis of transgenes expression relative to the *RplpO* mRNA in different brain areas and tissues of 6 months old control (n=3) or bigenic CMV-cre/full-length (n=3) or mutant (n=3) *FMR1* 5'UTR transgenic mice. (**C**) Immunohistochemistry against FMRpolyG N-terminus (8FM) of cerebellum and hippocampus areas of 6 months old bigenic CMV-cre/full-length or mutant *FMR1* 5'UTR transgenic mice. Scale bars, 10 μm. Sections were counterstained with Nissl staining. (**D**) Immunofluorescence against FMRpolyG N-terminus (green, 8FM antibody) and ubiquitin (red) on cerebellum areas of 6 months old bigenic CMV-cre/full-length or mutant *FMR1* 5'UTR transgenic mice. Scale bars, 10 μm. Nuclei were counterstained with DAPI. (**E**) Rotarod test. Time before falling from a rotating rod of 3 months old control (n=8) or bigenic CMV-cre/full-length (n=9) or mutant (n=9) *FMR1* 5'UTR transgenic male mice. (**F**) String test. Time to gain hindlimb traction for forelimb-hanging 3 months old control (n=8) or bigenic CMV-cre/full-length (n=9) or mutant (n=9) *FMR1* 5'UTR transgenic male mice. (**G**) Grip test. Maximal force relative to mouse body weight exerted to releases 3 months old control (n=8) or bigenic CMV-cre/full-length (n=9) or mutant (n=9) *FMR1* 5'UTR transgenic male mice holding a grid with their forepaws. (**H**) Open field. Numbers of rears during 5 minutes observation in open field of 3 months old control (n=8) or bigenic CMV-cre/full-length (n=9) or mutant (n=9) *FMR1* 5'UTR transgenic male mice. (**I**) Body weight of 2, 4 and 6 months old control (n=6) or bigenic CMV-cre/full-length (n=6) or mutant (n=6) *FMR1* 5'UTR transgenic male mice. (**J**) Left panel, immunofluorescence labeling of calbindin (red) of cerebellum

sections of 9 months old bigenic CMV-cre/full-length or mutant *FMR1* 5'UTR transgenic mice. Scale bars, 10 μm. Nuclei were counterstained with DAPI. Right panel, quantification of Purkinje cells (n=100) in cerebellum sections of 9 months old bigenic CMV-cre/full-length (n=3) or mutant *FMR1* 5'UTR (n=3) transgenic mice. (**K**) Kaplan-Meier survival curve of control (n=15), bigenic CMV-cre/full-length (n=15) or mutant (n=15) *FMR1* 5'UTR male and female transgenic mice. Error bars indicate s.e.m. Student t-test, * indicates $p<0.05$, ** indicates $p<0.01$ and *** indicates $p<0.001$.

Figure 4. Neuronal expression of FMRpolyG is pathogenic in mice.

(**A**) Schemes of the mouse transgene constructs. The human full-length 5'UTR of *FMR1* is represented as a red box, GFP ORF as a green box, loxP sites as blue boxes, while promoters, polyadenylation sequences and start and stop of translation are indicated by arrow or black lines. (**B**) Quantitative RT-PCR analysis of transgenes expression relative to the *RplpO* mRNA in different tissues of 6 months old control (n=3) or bigenic CMV-cre/full-length *FMR1* 5'UTR (n=3) or bigenic Nestin-cre full-length *FMR1* 5'UTR (n=3) mice. (**C**) Immunohistochemistry against FMRpolyG N- or C-terminal part (8FM or 9FM antibody) in the cerebellum, hippocampus and hypothalamus of 6 months old bigenic Nestin-cre/full-length *FMR1* 5'UTR mice. Scale bars, 10 μm. Sections were counterstained with H&E staining. (**D**) Rotarod test, time before falling of 3 months old control (n=6) or bigenic Nestin-cre/full-length *FMR1* 5'UTR (n=6) male mice. (**E**) Left panel, immunohistochemistry against calbindin of cerebellum sections of 10 months old control or bigenic Nestin-cre/full-length *FMR1* 5'UTR mice. Scale bars, 10 μm. Sections were counterstained with H&E staining. Right panel, quantification of Purkinje cells (n=50) in cerebellum sections of 10 months old control (n=3) or bigenic Nestin-cre/full-length *FMR1* 5'UTR (n=3) mice. (**F**) Immunohistochemistry against Gfap of hippocampal sections of 10 months old control or bigenic Nestin-cre/full-length *FMR1* 5'UTR mice. Scale bars, 50 μm. Sections were counterstained with H&E staining. (**G**) Representative image of 8 months old control or bigenic Nestin-cre/full-length *FMR1* 5'UTR mice. (**H**) Kaplan-Meier survival curves of control (n=10) or bigenic Nestin-cre/full-length *FMR1* 5'UTR (n=10) male and female mice. Error bars indicate s.e.m. Student t-test, * indicates $p<0.05$, ** indicates $p<0.01$.

Figure 5. The C-terminal part of FMRpolyG contributes to its toxicity.

(**A**) Immunofluorescence against FMRpolyG N-terminal part (8FM antibody, green) and nuclear lamina (Lmnb1, red) in primary cultures of E18 mouse cortical neurons transfected for the indicated time period with expanded CGG repeats embedded within the 5'UTR of *FMR1* and fused to the GFP in the glycine frame. Scale bars, 10 µm. Nuclei were counterstained with DAPI. (**B**) Left panel, representative images of primary cultures of E18 mouse cortical neurons transfected with GFP or FMRpolyG-GFP full-length or mutants. Scale bars, 10 µm. Nuclei were counterstained with DAPI. Right panel, schemes of the mutant constructs of FMRpolyG-GFP. The N-ter construct corresponds to the N-terminal part of FMRpolyG including its polyglycine repeats fused to the GFP. The C-ter construct corresponds to the last 42 amino acids of FMRpolyG fused to the GFP. Lower panel, quantification of neuronal cell viability of GFP-positive (n=100 cells, 3 independent transfections) transfected E18 mouse cortical neurons. (**C**) Progeny eclosion ratio (n=100, 3 independent crosses) of *Drosophila* expressing FMRpolyG full-length or deleted of its C-terminus compared to control driver line (*Actin5C-Gal4/+*). (**D**) Kaplan-Meier survival curve of *Drosophila* expressing FMRpolyG full-length or deleted of its C-terminus compared to control driver line (*Tub5-Gal4/+*). Error bars indicate s.e.m. Student t-test, * indicates $p< 0.05$, *** indicates $p< 0.001$.

Figure 6. FMRpolyG interacts with LAP2ß and alters its nuclear localization.

(**A**) Silver staining of proteins captured through consecutive anti-FLAG and anti-HA affinity purification steps from N2A neuronal cells transfected with FLAG-HA-tagged FMRpolyG full-length or deleted of its C-terminus. (**B**) Immunoblotting against endogenous Lap2ß protein of tandem-tag purified proteins from N2A cells expressing FLAG-HA-tagged FMRpolyG full-length or deleted of its C-terminus. (**C**) Immunoblotting against the HA or GFP tags of HA-tagged immunoprecipitated proteins from N2A cells transfected with HA-LAP2ß and FMRpolyG-GFP full-length or deleted of its N- or C-terminus. (**D**) Left panel, GFP fluorescence and immunofluorescence against endogenous Lap2ß (red) in primary cultures of E18 mouse cortical neurons transfected with FMRpolyG-GFP full-length or deleted of its C-terminus. Nuclei were counterstained with DAPI. Right panel, quantification of co-localization of Lap2ß with FMRpolyG-GFP in transfected E18 mouse cortical neurons (n=100 neurons, 3 independent transfections). (**E**) Immunohistochemistry against Lap2ß of cerebellum, hippocampal and hypothalamic areas of 9 months old bigenic CMV-cre/full-length or mutant *FMR1* 5'UTR transgenic

mice. Sections were counterstained with H&E staining. (**F**) Left panel, immunofluorescence against FMRpolyG N-terminus (green, 8FM antibody) and Lap2ß (red) on hippocampal areas of 9 months old bigenic CMV-cre/full-length or mutant *FMR1* 5'UTR transgenic mice. Nuclei were counterstained with DAPI. Right panel, quantification of co-localization of Lap2ß with FMRpolyG in bigenic CMV-cre/full-length or mutant *FMR1* 5'UTR transgenic mice (n=50 neurons, 3 mice). (**G**) Immunohistochemistry against LAP2ß of cerebellum areas of FXTAS individual or age-matched control. Sections were counterstained with H&E staining. (**H**) Immunofluorescence against the N-terminal part of FMRpolyG (green, 8FM antibody) and LAP2ß (red) on brain sections (hippocampal area) of FXTAS patients or age-matched controls. Nuclei were counterstained with DAPI. Scale bars, 10 μm. Error bars indicate s.e.m. Student t-test, *** indicates p< 0.001.

Figure 7. LAP2ß rescues neuronal cell death induced by FMRpolyG.

(**A**) Upper panel, immunofluorescence against the N-terminal part of FMRpolyG (green, 8FM antibody) and LAP2ß (red) on neuronal cultures differentiated 40 days from iPS cells of FXTAs patients or control individuals. Lower panel, quantification of LAP2ß co-localization with FMRpolyG in neurons from iPSC of FXTAS and control individuals (n=100 neurons, 3 independent cultures). (**B**) Upper panel, immunofluorescence against the N-terminal part of FMRpolyG (green, 8FM antibody) and the lamin B1 (LMNB1, red) on neuronal cultures differentiated 40 days from iPS cells of FXTAs patients or control individuals. Lower panel, quantification of lamin B1 alteration in FMRpolyG positive cells in neurons from iPSC of FXTAS and control individuals (n=100 neurons, 3 independent cultures). (**C**) Cell viability of neuronal N2A cells transfected (n=3 transfections) with FMRpolyG-GFP full-length or deleted of its N- or C-terminus and with a plasmid expressing RFP as control or Ha-tagged LAP2ß. Error bars indicate s.e.m. Student t-test, *** indicates p<0.001.

Figure 8. Tentative model of pathogenicity in FXTAS.

Expanded CGG repeats are translated into a polyglycine containing protein, FMRpolyG, through initiation to a non-canonical ACG codon located upstream to the CGG repeats. In FXTAS, higher expression of *FMR1* mRNA as well as increased translation and stability of

the expanded CGG repeats result in accumulation of FMRpolyG in nuclear aggregates that sequester the LAP2ß protein and alter the nuclear lamina architecture.

LEGENDS OF THE SUPPLEMENTARY FIGURES:

Supplemental table 1. Proteins interacting with FMRpolyG.

Proteins associated with HA-FLAG-tagged FMRpolyG expressed in mouse N2A neuronal cells were captured through consecutive anti-FLAG and anti-HA affinity purification steps and identified by orbitrap ion trap mass analyzer.

Supplemental videos 1 and 2 (related to figure 3).

Video recording of ledge test of 3 months old bigenic CMV-cre/full-length or mutant *FMR1* 5'UTR transgenic mice indicates that FMRpolyG-expressing mice present some gait instability and increase foot slippage compared to CGG RNA only-expressing mice.

Supplemental videos 3 to 6 (related to figure 5).

24 hours recording of primary cultures of E18 mouse cortical neurons transfected with RFP and with control GFP or FMRpolyG-GFP full-length or deleted of its N- or C-terminus. Expression of either FMRpolyG-GFP or its C-terminus fused to the GFP leads to neuronal cell death. In contrast, expression of the polyglycine stretch deleted of FMRpolyG C-terminal part leads to aggregates formation with no overt cell toxicity.

Supplemental figure 1. Translation of CGG repeats requires a near-cognate codon.

(**A**) Partial sequence of the 5'UTR FMR1 CGG99x GFP (glycine frame) plasmid (Addgene #63091). CMV promoter and sv40 polyadenylation sequences are indicated in black uppercase at the start and the end of the sequence, respectively. The plasmid multiple cloning site sequence is in black lowercase with Fse1 and Xba1 cloning site underlined. The 5'UTR sequence of human *FMR1* is in bold black uppercase with expanded CGG repeats in red. The ACG near cognate codon is also indicated in red. The stop codon of FMRpolyG and the ATG codon of GFP were deleted and replaced by a glycine codon (gga). The GFP sequence is in green uppercase. (**B**) GFP fluorescence of HeLa cells transfected with expanded CGG repeats embedded or not in the 5'UTR of *FMR1* and fused in all three possible frames with the GFP deleted of its ATG. (**C**) Immunoblotting against the GFP of HeLa cell lysate transfected with expanded CGG repeats embedded within the 5'UTR of *FMR1* and fused to the GFP in the glycine frame and treated with increasing quantity of lysostaphin, a glycine endopeptidase. (**D**) Upper panel, sequence of the mutant *FMR1* 5'UTR including a LysC restriction site. Lower panel, LC-MS/MS

spectra of the N-terminal part of the immunoprecipitated and LysC digested protein translated from expanded CGG embedded in the LysC mutant 5'UTR of *FMR1*. (**E**) Sequence alignment of the 5'UTR of *FMR1* from human, chimpanzee, macaque, gibbon, galago, wild boar, dog, rat and mouse with Kozac consensus sequence. ACG translation start of FMRpolyG is indicated in red, the conserved nucleotides in the Kozac sequence are indicated in bold black.

Supplemental figure 2. Generation of antibodies specific to FMRpolyG.

(**A**) Upper panel, amino acid sequence of FMRpolyG. The N- and C-terminal peptide sequences used to generate 8FM and 9FM mouse monoclonal antibodies are underlined. Lower panel, immunoblotting validation of 8FM and 9FM antibodies directed respectively against the N- and C-terminal parts of FMRpolyG on lysate of HeLa cells transfected with expanded CGG repeats embedded in the 5'UTR of *FMR1* and fused to the GFP in the glycine frame. (**B**) Immunofluorescence against the C-terminal part of FMRpolyG (red, 9FM antibody) and ubiquitin (green) on brain sections (hippocampal area) of FXTAS patients compared to a age-matched control individual. (**C**) Upper panel, putative amino acid sequence of the C-terminal part of FMRpolyA. The peptide sequence used to generate the 5FM mouse monoclonal antibody is underlined. Lower panel, immunoblotting validation of the 5FM antibody on lysate of HeLa cells transfected with a construct expressing the GFP fused to the putative C-terminal part of FMRpolyA. (**D**) Immunofluorescence against the C-terminal part of FMRpolyA (red, 5FM antibody) and ubiquitin (green) on brain sections (hippocampal area) of FXTAS patients. Scale bars, 10 μm. Nuclei were counterstained with DAPI.

Supplemental figure 3. Expression of FMRpolyG is pathogenic in mice.

(**A**) RNA FISH against CGG RNA foci (red) coupled to immunofluorescence against FMRpolyG N-terminal part (8FM antibody, green) in brain of 6 months old bigenic CMV-cre/full-length or mutant *FMR1* 5'UTR transgenic mice. Nuclei were counterstained with DAPI. (**B**) Quantification of FMRpolyG protein aggregates and CGG RNA foci in 3 or 9 months old control (n=3), bigenic CMV-cre/full-length (n=3) or mutant (n=3) *FMR1* 5'UTR transgenic mice (n=100 cells). (**C**) Immunohistochemistry against FMRpolyG C-terminal part (9FM) of cerebellum, hippocampal and hypothalamic areas of 3, 6 or 9 months old bigenic CMV-cre/full-length or mutant *FMR1* 5'UTR transgenic mice.

Sections were counterstained with Nissl staining. (**D**) Quantification of FMRpolyG protein aggregates in control (n=3), bigenic CMV-cre/full-length (n=3) or mutant (n=3) *FMR1* 5'UTR transgenic mice. (**E**) Immunofluorescence against ubiquitin (red) and the C-terminal part of FMRpolyA (5FM, green) on hippocampal areas of 6 months old bigenic CMV-cre/full-length or mutant *FMR1* 5'UTR transgenic mice. Nuclei were counterstained with DAPI. (**F**) Hematoxylin and eosin staining of cerebellum of 9 months old bigenic CMV-cre/full-length or mutant *FMR1* 5'UTR transgenic mice. Arrowheads indicate Purkinje cells. (**G**) Immunohistochemistry against Iba1 and Gfap of the hippocampal areas of 9 months old bigenic CMV-cre/full-length or mutant *FMR1* 5'UTR transgenic mice. Sections were counterstained with H&E staining. Scale bars, 10 μm. Error bars indicate s.e.m. Student t-test, * indicates $p<0.05$, ** indicates $p<0.01$, *** indicates $p<0.001$.

Supplemental figure 4. The 5'UTR sequence of *FMR1* impairs formation of RNA foci.

(**A**) Schematic description of the expanded CGG repeats constructs embedded in the human 5'UTR sequence of *FMR1* or cloned without any *FMR1* sequence. The 99 CGG repeats embedded in their natural 5'UTR sequence is deposited at Addgene, #63089. The construct with 60 CGG repeats inserted into the multiple cloning site of pcDNA3.1 is described in Sellier et al, 2010. (**B**) Upper panel, RNA FISH against CGG RNA foci in cells transfected for 24 hours with expanded CGG repeats embedded or not in the 5'UTR of *FMR1* (untagged). Scale bars, 10 μm. Nuclei were counterstained with DAPI. Lower panel, quantification of cells (n=50) with CGG RNA foci (n=3 transfection). Error bars indicate s.e.m. Student t-test, *** indicates $p<0.001$. (**C**) RT-PCR analysis of nuclear and cytoplasmic fractions of neuronal cells transfected for 24 hours with expanded CGG repeats embedded or not in the 5'UTR of *FMR1* (untagged). Correct nuclear and cytoplasmic fractionation was controlled by RT-PCR for nuclear U6 snRNA and for cytoplasmic *RPLP0* mRNA.

Supplemental figure 5. The C-terminal part of FMRpolyG contributes to its toxicity.

(**A**) Immunofluorescence against the FLAG tag (green) and lamin B1 (Lmnb1, red) in primary cultures of E18 mouse cortical neurons transfected for the indicated time period with expanded CGG repeats embedded within the 5'UTR of *FMR1* and fused to a FLAG tag in the glycine frame. Scale bars, 10 μm. Nuclei were counterstained with DAPI.

(**B**) Immunofluorescence against the FLAG tag (green) and lamin B1 (Lmnb1, red) in HEK293 cells transfected for the indicated time period with expanded CGG repeats embedded within the 5'UTR of *FMR1* and fused to a FLAG tag in the glycine frame. Scale bars, 10 µm. Nuclei were counterstained with DAPI. (**C**) Immunoblotting against the GFP of the soluble and insoluble fractions of neuronal cells transfected with expanded CGG repeats embedded within the 5'UTR of *FMR1* and fused to the GFP in the glycine frame. (**D**) Cell viability of N2A mouse neuronal cultures transfected with FMRpolyG-GFP full-length or deleted of its N- or C-terminal part (5 independent transfections). (**E**) Cell viability of N2A mouse neuronal cultures transfected with FMRpolyG-FLAG full-length or deleted of its C-terminal part (3 independent transfections). (**F**) Transgene mRNA expression in independent *Drosophila* lines expressing either FMRpolyG-GFP full-length or deleted of its C-terminus. RNA levels are significantly different from one line to another by student t-test. However, this is independent of the construct as there is no significant difference of transgene expression between FMRpolyG-GFP and polyG ΔCter-GFP lines by ANOVA. FMRpolyG-GFP line #2 and polyG ΔCter-GFP line #2 correspond to the *Drosophila* lines analyzed in figure 5C and present similar transgene expression. (**G**) Progeny eclosion ratio (n=100, 3 independent crosses) of independent *Drosophila* lines expressing FMRpolyG full-length or deleted of its C-terminus compared to control driver line (*Actin5C-Gal4/+*). Fly lines expressing FMRpolyG-GFP present a decrease eclosion rate compared to *Drosophila* lines expressing polyG ΔCter-GFP (one way ANOVA, $P<0.01$). Error bars indicate s.e.m. Student t-test, * indicates $p<0.05$, ** indicates $p<0.01$, *** indicates $p<0.001$.

Supplementary figure 6. FMRpolyG alters LAP2ß and lamin B1 nuclear organization.
(**A**) Left panel, GFP fluorescence and immunofluorescence against the FLAG tag (green) and LAP2ß (red) in HEK293 cells transfected for 24 hours with GFP as control or with expanded CGG repeats embedded within the 5'UTR of *FMR1* and fused either to the GFP or to a FLAG tag in the glycine frame. Right panel, quantification of LAP2ß co-localization with FMRpolyG (n=50 cells, 3 independent transfections). (**B**) Immunohistochemistry against LAP2ß of hippocampal areas of FXTAS individual or age-matched control. Sections were counterstained with H&E staining. (**A**) Left panel, GFP fluorescence and immunofluorescence against the FLAG tag (green) and lamin B1 (LMNB1, red) in HEK293 cells transfected for 24 hours with GFP as control or with expanded CGG

repeats embedded within the 5'UTR of *FMR1* and fused either to the GFP or to a FLAG tag in the glycine frame. Right panel, quantification of lamin B1 alterations in FMRpolyG positive cells (n=50 cells, 3 independent transfections). (**D**) Left panel, GFP fluorescence and immunofluorescence against lamin B1 (LmnB1, red) in primary cultures of E18 mouse cortical neurons transfected with FMRpolyG-GFP full-length or deleted of its C-terminus. Right panel, quantification of lamin B1 alteration in GFP-positive transfected neurons (n=100 neurons, 3 independent transfections). (**E**) Immunofluorescence against the N-terminal part of FMRpolyG (green, 8FM antibody) and lamin B1 (LMNB1, red) on brain sections (hippocampal area) of FXTAS patients or age-matched control individual. Scale bars, 10 µm. Nuclei were counterstained with DAPI. Error bars indicate s.e.m. Student t-test, ** indicates $p<0.01$, *** indicates $p<0.001$.

Supplementary figure 7. Generation of an iPS cell model of FXAS.

(**A**) Immunofluorescence against MAP2 on neuronal cultures differentiated 40 days from iPS cells of FXTAS or control individuals. (**B**) *FMR1* mRNA expression relative to *RPLP0* in neuronal cultures (n=3 independent cultures) differentiated 40 days from iPS cells of FXTAS or control individuals. (**C**) Left panel, RNA FISH against expanded CGG repeats (red) coupled to immunofluorescence against FMRpolyG N-terminus (8FM antibody, green) in neurons differentiated 40 days from iPS cells originating from FXTAS or control individuals. Right panel, quantification of CGG nuclear RNA foci and FMRpolyG nuclear aggregates in neurons differentiated 40 days from iPS cells originating from FXTAS or control individuals (n=100 neurons, 3 independent cell cultures). (**C**) Cell viability of neuronal N2A cells transfected with FMRpolyG-FLAG full-length or deleted of its C-terminus and with a plasmid expressing RFP as control or Ha-tagged LAP2ß (n=3 transfections). Scale bars, 10 µm. Nuclei were counterstained with DAPI. Error bars indicate s.e.m. Student t-test, *** indicates $p<0.001$.

III- SCLEROSE LATERALE AMYOTROPHIQUE ET DEMENCE FRONTO-TEMPORALE

III-1- Introduction

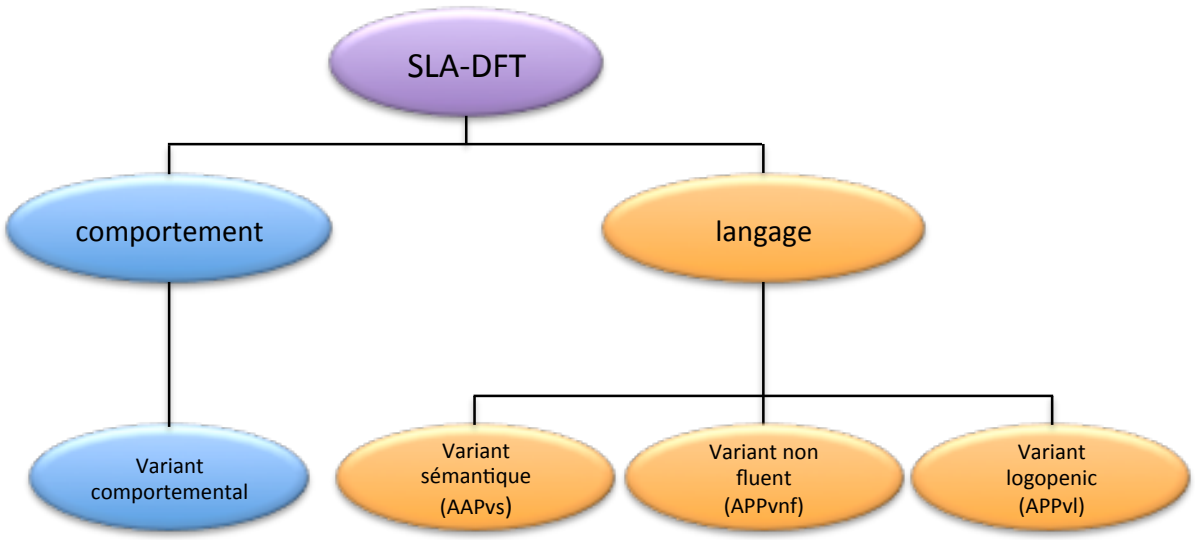
Autrefois considérées comme des entités cliniques séparées, il est désormais admis que la sclérose latérale amyotrophique (SLA) et la démence fronto-temporale (DFT) appartiennent à un même continuum. En effet, les patients atteints de SLA peuvent présenter des altérations cognitives et des troubles du comportement typiques de la DFT. De même, les patients atteints de DFT peuvent présenter des altérations motrices correspondant à une SLA (Goldstein and Abrahams, 2013). De plus, il est retrouvé des altérations histopathologiques communes, notamment des agrégats de protéines TDP43, dans ces deux maladies. Enfin, selon les patients, des mutations génétiques identiques, notamment du gène *C9ORF72*, peuvent donner soit des SLA, soit des DFT soit un mixte des deux entités (SLA/DFT). Il est toutefois à noter que ce continuum n'est pas parfait et qu'il existe des formes « pures » de SLA ou de DFT.

III-1-1 Présentation de la SLA/DFT et des facteurs génétiques impliqués.

III-1-1-a- La démence fronto-temporale (DFT)

La démence fronto-temporale recouvre un groupe hétérogène de présentations cliniques avec des phénotypes variables incluant un déficit de langage et/ou un déficit des fonctions cognitives et/ou des changements de comportement (Pressman *et al.*, 2014). La DFT représente la troisième forme de démence après Alzheimer et les démences à corps de Lewy (Vieira *et al.,* 2013). Sa prévalence est estimée à 10 à 15 cas pour 100 000 et environ 30 à 50% des cas sont d'origine génétique. La DFT résulte de la détérioration progressive des lobes frontaux et temporaux du cerveau. Ces désordres peuvent être génétiquement, cliniquement et pathologiquement hétérogènes. Ils sont divisés en deux catégories basées sur l'altération clinique majeure affectant soit le comportement soit le langage (figure 6, Zago *et al.*, 2011 *;* Sieben *et al*., 2012).

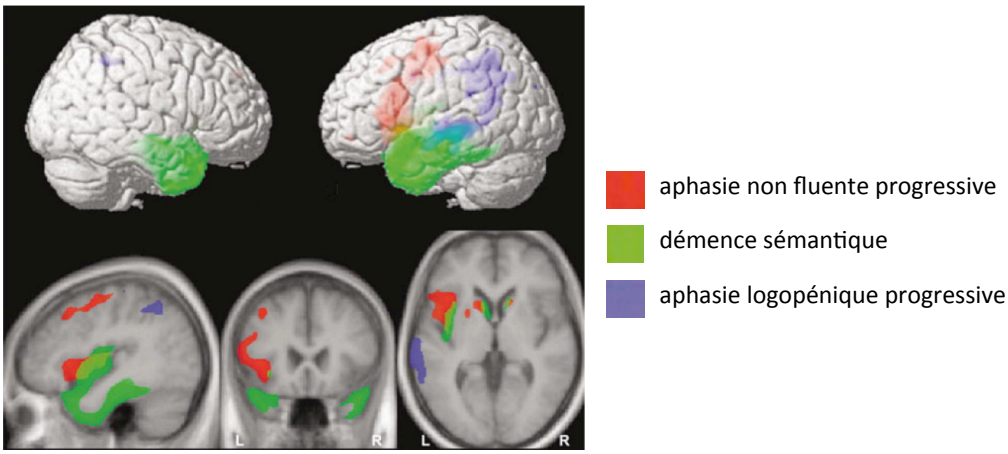
Figure 6 *: Classification des différentes sous-formes de DFT associées à la SLA (modifié d'après Zago et al., 2011)*



Le variant comportemental de la DFT est caractérisé par des changements du comportement et de la personnalité tels qu'une désinhibition, une apathie, une agressivité et/ou des difficultés à s'adapter à de nouvelles situations. Au stade initial de la maladie, la mémoire et les fonctions perceptives et spatiales sont préservées (Neary *et al.*, 1998 ; Josephs *et al.*; 2011). Le variant comportemental représente la moitié des formes de DFT et, généralement, les premiers signes de la maladie débutent avant l'âge de 65 ans avec une moyenne d'âge de début estimée à 58 ans (Josephs *et al.*, 2011).

L'aphasie progressive primaire (APP) est caractérisée par des difficultés de langage. Il en existe trois sous-variantes selon le type d'altération du langage. La variante sémantique (AAPvs) où le patient perd la conscience des significations de l'information non verbale, par exemple la connaissance d'objets connus, associée à une difficulté à nommer les objets. La variante non fluente (APPvnf) est caractérisée par une altération de la production du langage. L'élocution est ralentie, la parole est hésitante avec une articulation laborieuse mais le patient garde une compréhension des phrases simples et des mots. Et enfin, la variante logopénique (APPvl) qui est caractérisée par une production verbale diminuée et des difficultés à trouver les mots, la grammaire restant conservée (Baumann *et al.*, 2009). Ces différentes présentations cliniques reflètent la dégénérescence de zones différentes dans le cerveau (figure 7).

***Figure 7** : Régions atrophiques de différentes sous-formes d'aphasie progressive primaire comparativement à un groupe témoin (d'après Gorno-Tempini et al., 2004)*



Il est important de noter que ces altérations cliniques peuvent se chevaucher au cours de la progression de la maladie et que la distinction clinique entre les différents variants est compliquée voire impossible à des stades avancés de la maladie. La médiane de survie après l'apparition des premiers symptômes est de 6 à 11 ans indépendamment de l'âge de début de la maladie (Hodges *et al.*, 2003 ; Roberson *et al.*, 2005).

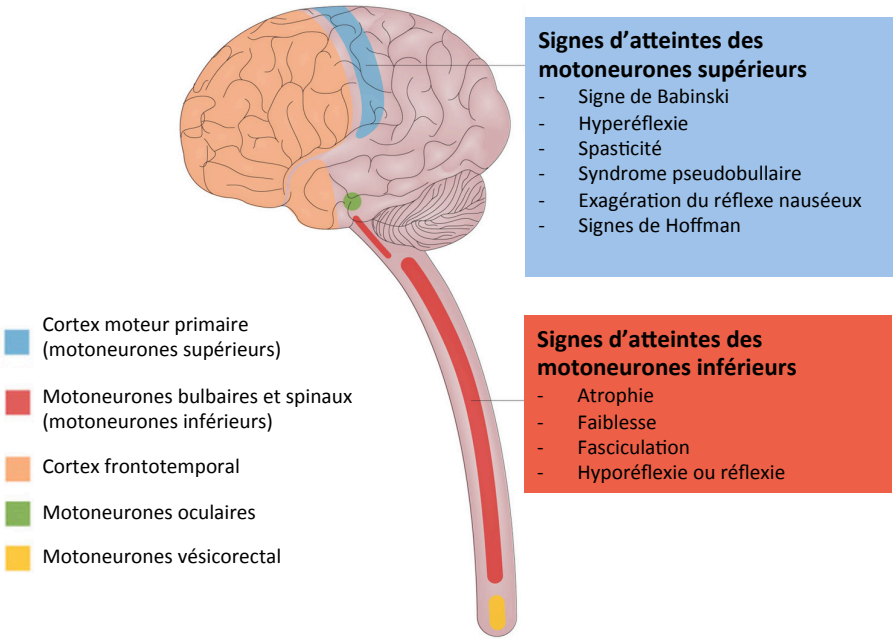
Enfin, une comorbidité de la SLA avec des altérations du comportement et des déficits cognitifs a été décrite dès le début du XX^{ème} siècle. Il apparaît aujourd'hui clairement qu'une DFT peut précéder, suivre ou coïncider avec le début de la SLA (Lomen-Hoerth *et al.*, 2003 ; Ringholz *et al.*, 2005 ; Phukan *et al.*, 2007 ; Van Langenhove *et al.*, 2012).

III-1-1-b- La sclérose latérale amyotrophique (SLA)

La sclérose latérale amyotrophique (SLA) également connue sous le nom de maladie de Charcot en Europe ou maladie de Lou Gehrig aux Etats Unis est la maladie la plus fréquente des maladies du motoneurone. L'incidence de la SLA en Europe est estimée à 2,16 pour 100 000 personnes par an (Logroscino *et al.*, 2010). Elle est caractérisée cliniquement par l'association à des degrés divers d'une paralysie progressive des muscles à innervation spinale et bulbaire, secondaire à une perte sélective et progressive des motoneurones centraux et périphériques dans le système nerveux central et de la corne antérieure de la moelle épinière (Wijesekera and Leigh, 2009 ; Swash and Desai, 2010).

Les principaux neurones touchés sont les motoneurones du cortex primaire moteur et les motoneurones bulbospinaux périphériques. Il est à noter que les neurones dans le cortex fronto-temporale sont aussi particulièrement touchés dans les cas de DFT/SLA. Enfin, les atteintes au niveau des neurones moteurs oculaires et vésico-rectaux sont rares et souvent observées dans le cas de maladie de longue durée d'évolution (figure 8, Swinnen and Robberecht, 2014).

***Figure 8** : Localisation des neurones affectés dans la SLA (modifié d'après Swinnen and Robberecht, 2014)*



La SLA peut se présenter sous deux formes principales : la forme bulbaire ou la forme spinale qui débute par l'atteinte d'un membre. Cette forme représente deux tiers des cas et débute généralement par une sensation de faiblesse accompagnée de fasciculations. Ces contractions s'accompagnent de crampes et d'une sensation de raideur dans les articulations et les membres qui rendent peu à peu les mouvements difficiles. Elle est due à la dégénérescence des motoneurones situés dans la moelle épinière et peut toucher le(s) membre(s) supérieur(s) ou inférieur(s). La forme bulbaire représente environ 30% des cas (Couratier *et al.*, 2000 ; Traynor *et al.*, 2010), les premiers symptômes sont des troubles de la phonation et de l'élocution se traduisant par une dysarthrie et une voix mal articulée. Le patient présente également des troubles de la déglutition. Ces deux formes peuvent se succéder ou se développer

simultanément, la maladie progressant presque toujours vers une forme complète (spinale et bulbaire).

La principale conséquence clinique liée à la dégénérescence des motoneurones est une faiblesse progressive des muscles et, au stade final de la maladie, une insuffisance respiratoire conduisant à la mort dans un délai de 36 à 60 mois après le début des symptômes (McGuire *et al.*, 1996 ; Pastula *et al.*, 2010 ; Al-Chalabi and Hardinam, 2013). Cependant, l'évolution de la maladie peut être très variable rendant difficile la prédiction de l'espérance de vie (Gordon *et al.*, 2012).

III-1-1-c- La SLA/DFT, un continuum entre deux maladies

Les recherches cliniques, en imagerie cérébrale et en neuropathologie, ont progressivement mis en évidence des chevauchements entre la SLA et la DFT. Dans la SLA, les lésions de la maladie s'étendent fréquemment au-delà du système moteur, incluant des structures cérébrales impliquées dans la cognition et les émotions. Plusieurs études retrouvent ainsi des troubles cognitifs ou comportementaux chez 50 % des patients atteints de SLA et parmi ceux-ci, 15 % répondent aux critères diagnostiques de DFT (Ringholz *et al.*, 2005 ; Murphy *et al.*, 2007). Inversement, 40% des patients atteints de DFT présentent des altérations des fonctions motrices dont 15% correspondent à une SLA (Lipton *et al.*, 2004 ; Burrell *et al.*, 2011).

De même, il existe un chevauchement entre SLA et DFT au niveau histopathologique. En effet, la plupart des cas de SLA et de DFT sont caractérisés par la présence d'inclusions positives à l'ubiquitine dans les neurones (Leigh *et al.*, 1991). En 2006, il a été montré que le principal composant de ces inclusions est la protéine TDP-43, codée par le gène *TARDBP* (Neumann *et al.*, 2006 ; Arai *et al.*, 2006). La distribution topographique de ces inclusions dans le cerveau est corrélée aux régions atteintes, suggérant fortement qu'un même mécanisme pathologique touche les régions corticales et les motoneurones et ceci dans les formes sporadiques comme dans les formes familiales. Toutefois, il existe aussi des cas de SLA ou de DFT sans agrégats de TDP-43, notamment lorsqu'elles sont dues à des mutations dans les gènes *SOD1*, *FUS* ou *MAPT* conduisant, respectivement, à des agrégats de protéine SOD1, FUS ou Tau (codée par MAPT). De plus, dans les SLA/DFT dues aux expansions G_4C_2 dans le gène *C9ORF72*, il a été observé des agrégats de protéine p62/SQSTM1 associée aux protéines DPR codées par ces répétitions (Al-

Sarraj *et al.*, 2011 ; Mann *et al.*, 2013). Etonnamment, ces agrégats ne sont pas localisés avec les inclusions de protéine TDP-43.

Enfin, la découverte fondamentale que des mutations dans certains gènes dont *TARDBP*, *TBK1* et *C9ORF72* sont responsables à la fois de SLA, de DFT ou de SLA/DFT a achevé de démontrer un continuum entre ces deux maladies (Kabashi *et al.*, 2008 ; Gitcho *et al.*, 2008 ; Rutherford *et al.*, 2008 ; Sreedharan *et al.*, 2008 ; Van Deerlin *et al.*, 2008 ; Yokoseki *et al.*, 2008 ; DeJesus-Hernandez *et al.*, 2011 ; Renton *et al.*, 2011 ; Cirulli *et al.*, 2015 ; Freischmidt *et al.*, 2015). En effet, environ 10 % des SLA sont d'origine familiale et une histoire familiale de démence est notée chez 30 à 50 % des malades atteints de DFT. Avec le séquençage à haut débit, une trentaine de gènes ont été rapportés comme étant associés à la SLA/DFT. Cependant, certains de ces gènes sont impliqués dans un faible nombre de cas comme *CHCHD10* (Johnson *et al.*, 2014 (a) ; Bannwarth *et al.*, 2014) ou le gène codant la matrine-3 (Johnson *et al.*, 2014 (b) ; Millecamps *et al.*, 2014). A l'inverse, d'autres gènes sont considérés comme « majeurs » car retrouvés mutés chez un plus grand nombre de patients. Ainsi des mutations dans les gènes *C9ORF72*, *SOD1*, *TBK1*, *TARDBP* et *FUS/TLS* sont responsables de 50% des cas de SLA familiales. De même, des mutations dans les gènes *MAPT*, progranuline (*GRN*) et *C9ORF72* sont responsables de 60 à 70% des cas de DFT familiales (Baker *et al.*, 2006 ; Cruts *et al.,* 2006 ; DeJesus-Hernandez *et al.*, 2011; Renton *et al.*, 2011).

La découverte de différents gènes mutés dans la SLA/DFT conforte donc l'idée d'un continuum entre ces deux maladies. Toutefois, il est important de noter qu'il existe aux extrémités de ce continuum des cas de SLA et des DFT « pures ». Ainsi, des mutations dans le gène *SOD1* sont uniquement retrouvées dans la SLA (Rosen *et al.*, 1993). De même, des mutations dans les gènes *MAPT* ou *GRN* sont spécifiquement associées à la DFT (Clark *et al.*, 1998 ; Baker *et al.,* 2006 ; Cruts *et al.*, 2006)

III-1-1-d Les modificateurs génétiques de la SLA/DFT

Bien que la connaissance des bases génétiques de la SLA/DFT ne cesse de progresser et permet ainsi de mieux comprendre les mécanismes altérés dans la maladie, les différences dans la pénétrance et la sévérité de la maladie pour une même altération génétique restent mal comprises. En effet, l'âge de début de la maladie ou du temps de survie après le début de la maladie sont variables entre deux individus d'une même famille présentant une même mutation

à l'origine de la SLA/DFT. Ces observations suggèrent l'existence de facteurs modificateurs.

En effet, des variants génétiques peuvent, en association avec d'autres facteurs génétiques ou environnementaux, moduler l'apparition et l'évolution de la maladie. Par exemple, le polymorphisme p.A152T dans le gène codant la protéine *MAPT* augmente le risque de développer une DFT (Coppola *et al.*, 2012 ; Lee *et al.*, 2013 (a)). De même, un polymorphisme dans le gène *TMEM106B* modifie la sévérité et l'âge d'apparition des DFT dues à des mutations de la progranuline (van Blitterswijk *et al.*, 2014). Enfin, le polymorphisme rs12608932 dans le gène *UNC13A* influence la susceptibilité de développer une SLA et est associé à une médiane de survie plus courte (van Es *et al.*, 2009 ; Diekstra *et al.*, 2012).

L'ataxine-2 (ATXN2) est également largement décrite comme un modificateur dans la SLA/DFT (Elden *et al.*, 2010 ; Lee *et al.*, 2011 ; Van Damme *et al.*, 2011 ; Lattante *et al.*, 2014). Ce gène code pour une protéine cytoplasmique de 140 kDa impliquée dans la stabilité des ARNm. L'ataxine-2 contient dans sa région N-terminale une expansion de polyglutamines (polyQ) comprenant le plus souvent 22-23 polyQ (Ross *et al.*, 2011). Cette expansion de polyQ, codée par des codons CAG, est généralement interrompue par des codons CAA, codant également des glutamines, pour former la séquence (CAG)8-CAA-(CAG)4-CAA-(CAG)8 (Imbert *et al.*, 1996). Ces répétitions CAG sont instables et peuvent s'étendre conduisant alors à des expansions pathogènes de polyQ. Ainsi, des expansions ininterrompues de plus de 37 CAG conduisent à l'ataxie spinocérébelleuse de type 2 (SCA2) caractérisée par une dégénérescence des cellules de Purkinje du cervelet associée avec le temps à des altérations neuronales plus larges (Imbert *et al.*, 1996 ; Pulst *et al.*, 1996 ; Sanpei *et al.*, 1996).

De plus, des expansions d'environ 40 CAG semblent augmenter le risque de développer une paralysie supranucléaire progressive (PSP) ou une maladie de parkinson, notamment dans les populations d'origine asiatique (tableau 1). Au contraire, dans les populations européennes, une taille intermédiaire de CAG (27 à 33 polyQ) est associée de manière significative avec la SLA (Elden *et al.*, 2010 ; Lee *et al.*, 2011 ; Van Damme *et al.*, 2011). Au point de vue histopathologique, environ 27% des patients atteints de SLA (avec une taille de polyQ normale ou intermédiaire) présentent des agrégats d'ataxine-2 dans les motoneurones alors que moins de 5% des personnes contrôles en présentent (Elden *et al.*, 2010). De plus, les patients portant une taille intermédiaire de polyQ dans l'ataxine-2 présentent des agrégats de TDP-43 avec une forme filamenteuse alors que les patients atteints de SLA avec une taille de polyQ normale

présentent des inclusions de TDP-43 plutôt de forme sphérique (Hart *et al.*, 2012). Enfin, la présence d'une taille intermédiaire de polyQ dans l'ataxine-2 est un modificateur fort du risque de développer une SLA chez les patients porteur d'une expansion G_4C_2 dans le gène *C9ORF72* (Lattante *et al.*, 2014).

Ces résultats suggèrent que l'ataxine-2 avec une expansion intermédiaire de polyQ joue un rôle important dans la pathogenèse de la SLA.

***Tableau 1** : Expansion polyQ dans l'ataxine-2 et syndromes neurodégénératifs (modifié d'après van den Heuvel et al., 2014)*

Longueur des répétitions polyQ dans l'ataxine-2				
maladie	longueur des répétitions	séquence		références
normal	14-31	interrompue	Le plus commun 22 répétitions : (CAG)8CAA(CAG)4CAA(CAG)8	Imbert *et al.*, 1996
SLA	24-39	interrompue	Valeur seuil variable selon les populations : ≥27 cohorte nord américaine ≥29 cohorte européenne	Elden *et al.*, 2010 Lee *et al.*, 2011 Van Damme *et al.*, 2011
SCA2	>31	répétitions CAG pures	31-33 répétitions associées avec un début tardif de la maladie	Imbert *et al.*, 1996 Fernandez *et al.*, 2011
parkinsonisme	31-44	interrompue	Association trouvée dans des cohortes chinoises ou taiwanaise	Lu *et al.*, 2004 Shan *et al.*, 2001
paralysie supranucléaire progressive	31-33	interrompue		Ross *et al.*, 2011

III-1-2- SLA/DFT liée à une mutation dans le gène *C9ORF72*

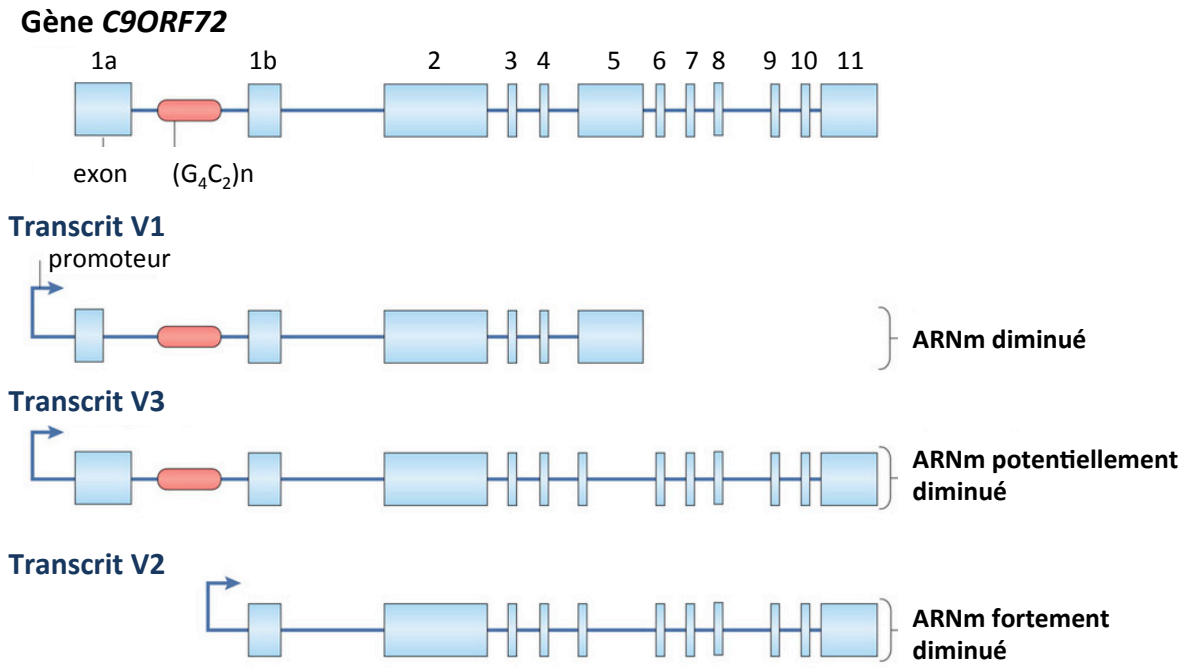
En 2011, la cause génétique la plus fréquente de SLA et/ou DFT a été identifiée comme étant une expansion d'hexanucléotides G_4C_2 localisée entre les 2 premiers exons non codant du gène *C9ORF72* (DeJesus-Hernandez *et al.*, 2011 ; Renton *et al.*, 2011).

III-1-2-a- C9ORF72 : gène et protéine

Le gène *C9ORF72* contient deux promoteurs possibles (exons 1a et 1b) et 11 exons dont un intron potentiellement retenu de façon alternative (intron alternatif 6) conduisant à une isoforme protéique dite courte de 222 acides aminés ou à l'isoforme longue de 481 acides aminés (figure 9). La protéine C9ORF72 (forme longue) est conservée à travers le règne des métazoaires, excepté chez les insectes, avec un haut degré d'homologie au niveau de la

séquence nucléotidique et de la séquence en acides aminés entre l'homme, la souris, le rat et le poulet (Xiao *et al.*, 2016).

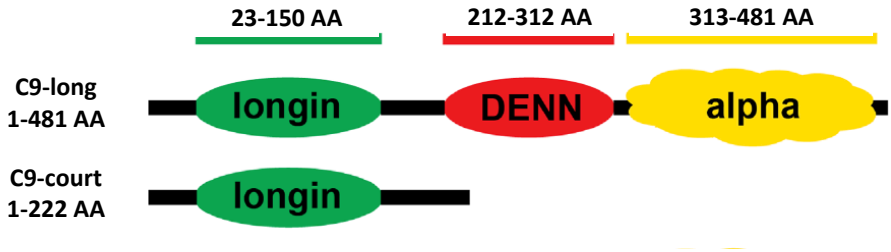
***Figure 9** : Schéma du gène C9ORF72 et des différents transcrits obtenus après épissage alternatif (modifié d'après Haeusler et al.,2016).*



Des études menées chez l'homme, la souris et le zébrafish montrent une expression élevée du gène *C9ORF72* dans le cerveau, la moelle épinière et notamment dans les neurones (Ciura *et al.*, 2013 ; Suzuki *et al.*, 2013 ; Kim *et al.*, 2014). Des analyses immunohistochimiques ont révélé que la forme courte de la protéine C9ORF72 est localisée au niveau de la membrane nucléaire alors que la forme longue est détectée dans le cytoplasme (Xiao *et al.*, 2015). Des analyses bioinformatiques suggèrent que la protéine C9ORF72 contient des domaines longin et DENN (differentially expressed in normal and neoplastic cell) (figure 10 ; Zhang *et al.*, 2012 ; Levine *et al.*, 2013) typiques des facteurs d'échange nucléotidique GDP-GTP (GEF) pour les protéines Rab GTPases, qui sont des régulateurs majeurs de la biogenèse, du transport et de l'adressage de vésicules aux membranes (Marat *et al.*, 2011). En accord avec un rôle de C9ORF72 comme un facteur d'échange nucléotidique GDP-GTP, cette protéine a été identifiée comme associée aux Rab GTPases 1, 5 et 7 et impliquée dans l'autophagie et/ ou l'endocytose (Farg *et*

al., 2014). Toutefois, les fonctions moléculaires et cellulaires de la protéine C9ORF72 (courte ou longue) restent largement incomprises.

***Figure 10** : Structure des isoformes de la protéine C9ORF72 (modifié d'après Xiao et al., 2016)*



III-1-2-b- Expansions G_4C_2 dans le gène *C9ORF72* et SLA/DFT

L'expansion de répétitions G_4C_2 dans le gène *C9ORF72* est une mutation majeure pour comprendre la SLA/DFT. En effet, cette mutation est identifiée dans 80% des formes de SLA/DFT familiales, 20 à 50% des SLA familiales, 10 à 20 % des DFT familiales, mais aussi dans 5 à 10% des SLA ou DFT sporadiques dans les populations originaires d'Europe et d'Amérique du Nord (Boeve *et al.,* 2012; Majounie *et al.,* 2012).

Le seuil où le nombre de répétitions G_4C_2 devient pathologique n'est pas clairement déterminé. Environ 90% des personnes saines présentent moins de 8 répétitions (Rutherford *et al.*, 2012), et il est supposé qu'au-delà de 30 répétitions, l'expansion G_4C_2 devient pathologique, toutefois avec des exceptions (Gami *et al.*, 2015 ; Xi *et al.*, 2015 ; Gijselinck *et al.*, 2016). La toxicité de ces expansions G_4C_2 augmente avec l'âge avec une pénétrance de 0% à 30 ans, de 50% à 58 ans et de plus de 90% à 80 ans (Majounie *et al.*, 2012 ; Benussi *et al.*, 2014). Toutefois, il existe des personnes porteuses d'expansions G_4C_2 non atteintes de SLA ou de DFT à l'âge de 80 ans (Galimberti *et al.*, 2013).

Outre la SLA/DFT, une expansion de répétitions G_4C_2 a aussi été identifiée comme une cause rare de phénocopie de la maladie de Huntington (Kostic *et al.*, 2014). De plus, les patients atteints de DFT ou SLA/DFT avec une expansion G_4C_2 présentent davantage de problèmes psychiatriques, notamment des hallucinations, que les patients avec d'autres mutations responsables de SLA/DFT (Arighi *et al.*, 2012 ; Dobson-Stone *et al.*, 2012 ; Snowden *et al.*, 2012).

Enfin, il est important de noter qu'il n'a pas été identifié de mutation responsable de SLA/DFT dans le gène *C9ORF72* autre que ces expansions de répétitions G_4C_2. Une mutation d'épissage (c.601-2A>G), probablement perte de fonction, a bien été récemment identifiée mais dans un seul patient SLA/DFT et son importance pour la pathologie est incomprise (Liu *et al.*, 2016). Etudier la toxicité de ces répétitions G_4C_2 est donc cruciale pour mieux comprendre la SLA/DFT.

III-1-2-c- Mécanismes pathologiques des expansions G_4C_2 dans *C9ORF72*

Les mécanismes par lesquels la présence d'une expansion de répétitions G_4C_2 dans le gène *C9ORF72* entraîne une neurodégénérescence ne sont pas clairement identifiés mais pourraient faire intervenir trois mécanismes différents non exclusifs (Figure 11).

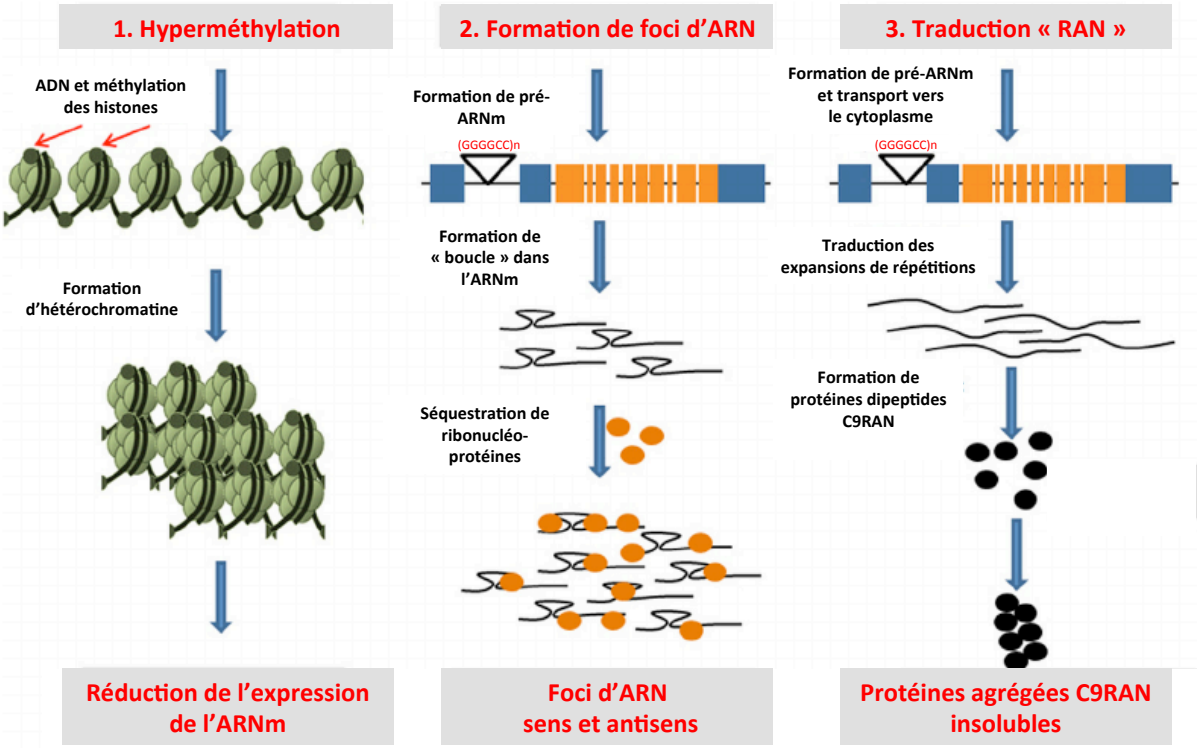
- Les ARN transcrits contenant les expansions de répétitions G_4C_2 peuvent lier différentes protéines de liaison à l'ARN, provoquant ainsi une perte de fonction de ces protéines. Il a ainsi été montré la présence d'agrégats nucléaires d'ARN G_4C_2 dans des échantillons de cerveaux et de moelle épinière et dans des cellules iPS (induced pluripotent stem cell) dérivées de patients SLA/DFT (de Jesus-Hernandez *et al.,* 2011 ; Donnelly *et al.,* 2013 ; Almeida *et al.*, 2013; Lagier-Tourenne *et al.,* 2013; Mizielinska *et al.*, 2013). Ces agrégats d'ARN sont positifs pour différentes protéines liant l'ARN comme la nucléoline, ADARB2, hNRNPA1, hnRNPH ou PURalpha (Donnelly *et al.,* 2013 ; Sareen *et al.,* 2013 ; Lee *et al.*, 2013 (b); Cooper-Knock *et al.*, 2014; Haeusler *et al.*, 2014 ; Prudencio *et al.*, 2015 ; Conlon *et al.*, 2016). Toutefois, et contrairement aux dystrophies myotoniques, la titration de ces protéines de liaison aux ARN semble faible et ne conduit qu'à des altérations subtiles de l'épissage alternatif dans les patients SLA/DFT (Prudencio *et al.*, 2015 ; Conlon *et al.*, 2016). Les conséquences pathologiques de la présence de ces protéines associées aux agrégats d'ARN G_4C_2 restent donc encore à déterminer.

- Les expansions de répétitions G_4C_2, mais aussi de répétitions antisens C_4G_2, peuvent être traduites en absence de codon d'initiation ATG par un mécanisme encore mal compris, et ainsi donner cinq protéines constituées de répétitions de dipeptides (DPR : glycine-alanine ; glycine-proline ; glycine-arginine ; proline-alanine et proline-arginine ; Ash *et al.*, 2013;

Gendron *et al.*, 2013; Mori *et al.*, 2013 (a et b); Zu *et al.*, 2013). La présence de ces protéines a été confirmée dans des modèles de cellules iPS de patients SLA/DFT, mais aussi dans des modèles de souris exprimant les répétitions G_4C_2 (Chew *et al.*, 2015; O'Rourke *et al.*, 2015; Peters *et al.*, 2015 ; Jiang *et al.*, 2016). Ces inclusions cytoplasmiques de DPR sont négatives pour la présence de TDP-43 mais positives pour la protéine p62/SQSTM1 (Al-Sarraj *et al.*, 2011 ; Troakes *et al.*, 2012 ; Mann *et al.*, 2013). Toutefois, et contrairement aux agrégats de TDP-43, ces inclusions de DPR sont majoritairement trouvées dans les régions du cerveau peu affectées par la pathologie (Baborie *et al.*, 2015 ; Gomez-Deza *et al.*, 2015 ; Mackenzie *et al.*, 2015 ; Davidson *et al.*, 2016). Enfin, l'expression de ces DPR sous la dépendance d'un codon ATG est toxique en transfection cellulaire et chez la drosophile, notamment en interférant avec le transport nucléocytoplasmique (Kwon *et al.*, 2014; Mizielinska *et al.*, 2014; Wen *et al.*, 2014; Zhang *et al.*, 2014 ; Jovicic *et al.*, 2015; Tao *et al.*, 2015 ; Tran *et al.*, 2015 ; Boeynaems *et al.,* 2016).

- Tout comme dans le cas du syndrome de l'X fragile, des modifications épigénétiques, notamment l'hyperméthylation d'îlots CpG en amont des expansions de répétitions G_4C_2 entraînent une diminution de la transcription du gène *C9ORF72* (Xi *et al.,* 2013; Belzil *et al.,* 2014). Différentes études ont ainsi confirmé une diminution de l'expression de l'ARNm de *C9ORF72* dans le sang, le cerveau et la moelle épinière de patients portant des expansions G_4C_2 dans le gène *C9ORF72* (DeJesus-Hernandez *et al.*, 2011; Gijselinck *et al.*, 2012; Almeida *et al.*, 2013 ; Fratta *et al.,* 2013 ; Belzil *et al.,* 2013 ; Donnelly *et al.,* 2013 ; Davidson *et al.,* 2014). Une étude plus complète, menée sur 56 patients avec une expansion G_4C_2, a non seulement confirmé cette réduction d'expression mais a aussi identifié une corrélation entre le niveau d'expression de *C9ORF72* et la survie des patients (van Blitterswijk *et al.*, 2015). Cette diminution des transcrits *C9ORF72* a été confirmée au niveau protéique (Waite *et al.,* 2014; Xiao *et al.,* 2015). Enfin, la perte d'expression de *C9ORF72* chez le poisson zèbre conduit à un phénotype locomoteur avec une dégénérescence des motoneurones (Ciura *et al.*, 2013). Ces résultats suggèrent donc qu'une haploinsuffisance de *C9ORF72* puisse contribuer à la SLA/DFT.

Figure 11 *: Schéma représentant les trois mécanismes toxiques impliqués lors de la présence d'expansions G_4C_2 dans la gène C9ORF72 (modifié d'après Vats et al., 2014)*



Il apparaît donc que les mécanismes responsables de la toxicité des expansions G_4C_2 sont complexes et probablement multiples. En effet, si l'expression de protéines DPR à partir des répétitions G_4C_2 est bien établie, les régions d'accumulation de ces inclusions corrèlent peu ou mal avec la pathologie. De même, si la diminution d'expression de C9ORF72 chez les patients porteurs d'expansion G_4C_2 est confirmée par de multiples études, les conséquences pathologiques de cette diminution sont encore mal comprises, et ce d'autant plus que les fonctions de la protéine C9ORF72 sont peu connues (Farg *et al.*, 2014). Afin de mieux comprendre l'importance de C9ORF72 dans la SLA/DFT, je me suis donc intéressée aux fonctions moléculaires et cellulaires de cette protéine. Ayant identifié un rôle de C9ORF72 dans l'autophagie (cf. Partie résultats), je vais décrire ci-dessous succinctement ce mécanisme et son importance pour la SLA/DFT.

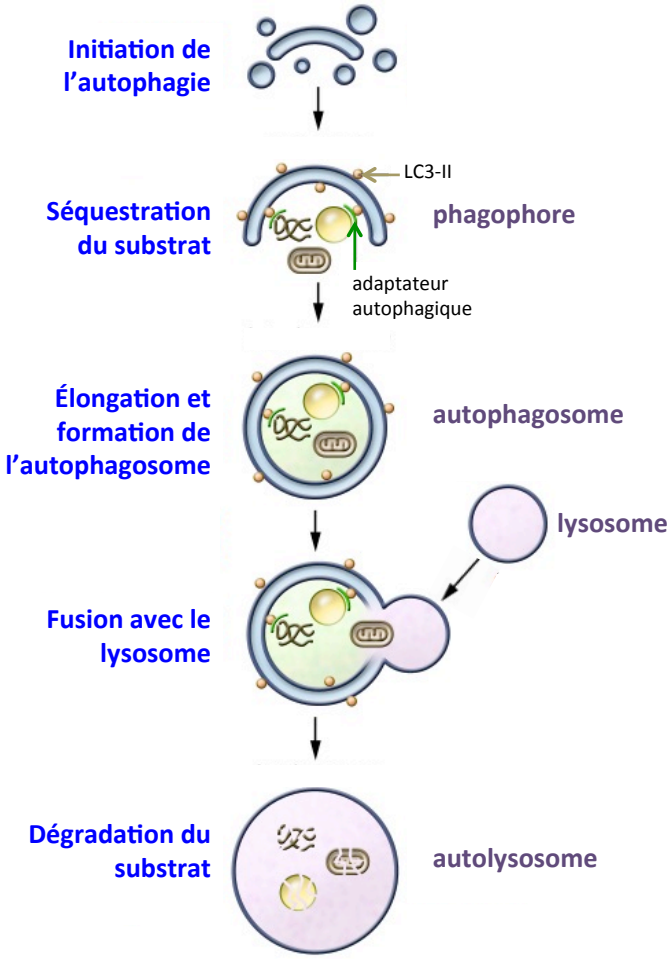
III-1-3- Rôle de l'autophagie dans la SLA/DFT

III-1-3-a- Généralités sur le mécanisme de l'autophagie

L'autophagie fait partie avec la dégradation par le protéasome des deux mécanismes essentiels qui permettent aux cellules d'éliminer les protéines anormales ou présentes en quantité excessive (Klionsky and Emr, 2000 ; Rubinsztein, 2006). Trois types d'autophagie sont connus à ce jour (Mizushima *et al.*, 2007). La microautophagie est un mécanisme encore peu étudié où les lysosomes engloutissent directement et de façon non spécifique le cytoplasme les environnant (Li *et al.*, 2012). Au contraire, l'autophagie régulée par les protéines chaperonnes (CMA) est caractérisée par sa sélectivité. Dans la CMA, seules les protéines possédant une séquence consensus de type KFERQ sont reconnues par la protéine LAMP2 (lysosome-associated membrane protein type 2) et transférées dans la lumière du lysosome pour être dégradées (Dice *et al.*, 2007). Enfin, la macroautophagie (ou autophagie) repose sur la formation d'une structure à double membrane, l'autophagosome, autour des éléments à dégrader. L'autophagosome fusionne ensuite avec le lysosome pour former l'autolysosome en vue de dégrader les composants absorbés par l'autophagosome (Klionsky, 2007). La dégradation *via* l'autophagie assure la régulation et l'élimination d'agrégats de protéines, de bactéries intracellulaires ou d'organites altérés qui ne peuvent être dirigés vers le protéasome du fait de leur taille.

Le processus autophagique est arbitrairement divisé en cinq étapes : l'initiation, l'élongation, la maturation, la fusion et la dégradation (figure 12). L'initiation débute par la formation d'une double membrane au point de nucléation appelée phagophore. Ensuite, cette membrane s'allonge pour séquestrer une partie du contenu cytoplasmique et former une vésicule délimitée par une double membrane : l'autophagosome. Lors de la maturation, l'autophagosome fusionne avec un lysosome pour former l'autolysosome. Et enfin, lors de la dégradation, le contenu de l'autophagosome est digéré par les hydrolases acides lysosomales (Mizushima, 2007)

Figure 12 *: Schéma représentant les différentes étapes du processus autophagique (modifié d'après Frake et al., 2015).*

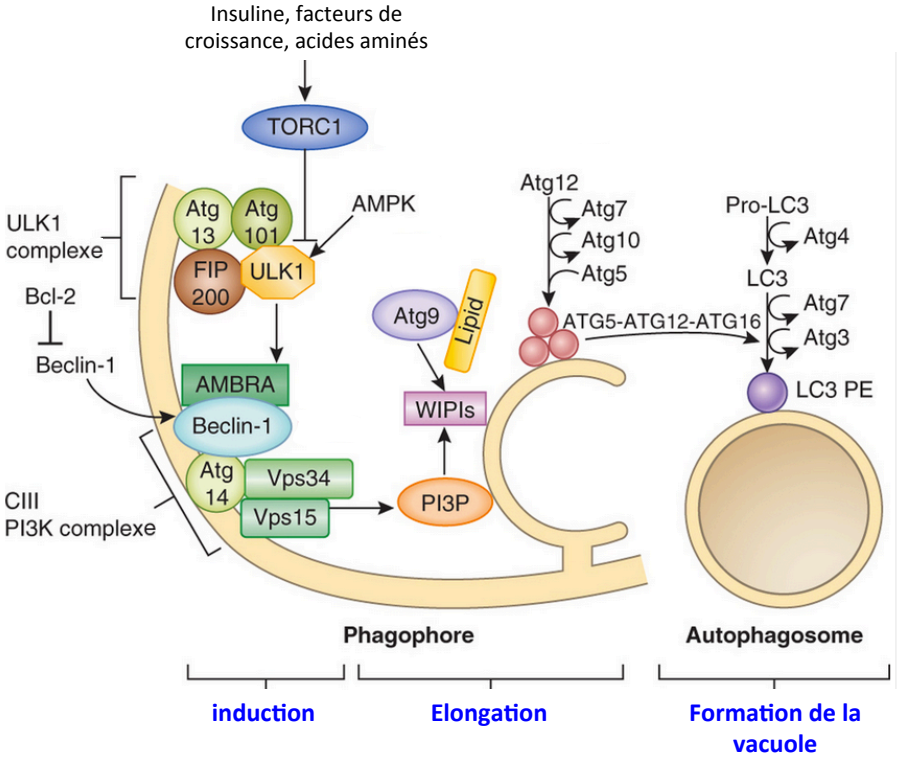


Deux mécanismes sont connus comme activant l'autophagie, l'absence de nutriments et d'énergie (autophagie non spécifique), et la poly-ubiquitinylation en K63 des composants à dégrader (autophagie spécifique).

En cas de privation de glucose ou d'acides aminés, la kinase AMPK phosphoryle et inhibe le complexe mTOR, dont l'une des nombreuses fonctions est d'inhiber par phosphorylation le complexe kinase ULK1/2-ATG13-ATG101-FIP200. De plus, AMPK phosphoryle et active directement ce complexe ULK1/2 (Ganley *et al.*, 2009). L'activation de ce dernier entraîne la phosphorylation de la Beclin-1, qui forme avec les protéines ATG14, VSP34 et VSP15 le complexe PI3K CIII chargé de la phosphorylation des phosphoinositides en Ptdins3P au niveau de la structure pré-autophagosomale (Simonsen and Tooze, 2009). Ces phosphoinositides 3

phosphates sont reconnus par la protéine WIPI2 qui va déclencher la formation et l'élongation de l'autophagosome. En effet, WIPI2 recrute la protéine ATG16L1 du complexe ATG16-ATG12-ATG5 (Dooley *et al.*, 2014), qui après l'action successive des protéines ATG4, ATG7 et ATG3 va induire la conjugaison des protéines LC3A ou LC3B (LC3-I) à un phosphatidyléthanolamine pour produire la forme LC3-II lipidée (LC3-II) (figure 13). Les protéines LC3A-II ou LC3B-II sont localisées et marquent la membrane de l'autophagosome (Hanada *et al.*, 2007 ; Nakatogawa *et al.*, 2007). Celui-ci est alors transporté, *via* les microtubules, vers le centre organisateur des microtubules où sont présents les lysosomes, afin de fusionner avec ces derniers pour former un autolysosome (Marambio *et al.*, 2010). Le contenu des lysosomes est dégradé par les protéases acides lysosomales et les acides aminés ou autres produits de dégradation sont réutilisés par la cellule.

Figure 13 *: Induction de l'autophagie et biogenèse de l'autophagosome (modifié d'après Nixon, 2013)*



L'autophagie contribue aussi à la dégradation sélective de différents substrats ubiquitinylés tels que les agrégats de protéines, les organelles dysfonctionnelles ou des microorganismes intracellulaires (Xie and Klionsky, 2007 ; Kirkin *et al.*, 2009). Cette spécificité requiert la présence de protéines adaptatrices, dites récepteurs autophagiques, qui se lient à la fois aux chaînes de polyubiquitine fixées en K63 aux protéines à dégrader et aux

autophagosomes *via* les protéines LC3-II (Filimonenko *et al.*, 2010 ; Rogov *et al.*, 2014). Il existe différents récepteurs autophagiques tels que p62/SQSTM1, NBR1, NDP52, optineurine (OPTN) qui sont exprimés sélectivement dans certains tissus et/ou qui reconnaissent spécifiquement les différents substrats à dégrader. Ainsi l'optineurine apparaît importante pour l'autophagie des mitochondries altérées (mitophagie), NDP52 serait plutôt spécifique des bactéries intracellulaires (xénophagie) et p62/SQSTM1 serait impliqué dans l'autophagie des agrégats protéiques (agrégophagie) (Johnson *et al.*, 2012). Tout comme la privation en énergie et en acides aminés active l'autophagie *via* le complexe kinase ULK1/2, l'autophagie sélective peut aussi être activée par phosphorylation, mais par la kinase TBK1 (TANK-binding kinase 1). En effet, l'optineurine qui interagit avec TBK1 peut recruter cette kinase au site de formation de l'autophagosome où TBK1 va phosphoryler l'optineurine et p62/SQSTM1 et renforcer ainsi leur liaison à l'ubiquitine et aux protéines LC3-II (Kachaner *et al* ; 2012 ; Maruyama and Kaxakami 2013). Ainsi, TBK1 stimule l'autophagie des protéines ubiquitinylées liées aux bactéries intracellulaires ou aux mitochondries altérées (Morton *et al.*, 2008 ; Gleason *et al.*, 2011 ; Pilli *et al.*, 2012 ; Heo *et al.*, 2015 ; Lazarou *et al.*, 2015 ; Matsumoto *et al.*, 2015). TBK1 est également important pour la maturation de l'autophagosome en autolysosome conduisant à la dégradation de p62/SQSTM1 et de ses protéines cargos (Pilli *et al.*, 2012).

III-1-3-b- L'autophagie dans la SLA/DFT.

Une des caractéristiques de la SLA/DFT est la présence d'inclusions cytoplasmiques dans le cerveau et les motoneurones des patients contenant notamment la protéine TDP-43 ubiquitinylée. TPD43 est une protéine encline à s'agréger et connue pour être éliminée par autophagie (Filimonenko *et al.*, 2007 ; Ju *et al.*, 2009 ; Wang *et al.*, 2012). De plus, des inclusions cytoplasmiques contenant LC3-II et p62/SQSTM1, des marqueurs caractéristiques des autophagosomes, ont été détectées dans les motoneurones de patients SLA sporadiques (Sasaki *et al.*, 2011). L'accumulation d'agrégats de TDP-43, p62/SQSTM1 et LC3 suggère donc que les mécanismes de dégradation des protéines, et notamment l'autophagie, sont altérés dans la SLA/DFT.

Par ailleurs, de nombreuses mutations dans des gènes impliqués dans les mécanismes de dégradation des protéines et l'autophagie sont impliquées dans la SLA/DFT. C'est notamment le cas de mutations dans les gènes codants pour l'ubiquilin-2 (*UBQLNL2*) ou VCP/p97, qui sont

impliqués dans la reconnaissance des protéines ubiquitinylées (Johnson *et al.*, 2010 ; Deng *et al.*, 2011 ; Gellera *et al.*, 2013). Des mutations dans le gène *CHMP2B* codant pour une sous-unité du complexe ESCRT-III impliquée dans la fermeture des vésicules notamment d'autophagie sont responsables de rares cas de DFT (Skibinski *et al.*, 2005 ; Parkinson *et al.*, 2006 ; Lindquist *et al.*, 2008). Des mutations dans les adaptateurs autophagiques tels que p62/SQSTM1 et l'optineurine sont également impliquées dans la SLA/DFT (Maruyama *et al.*, 2010 ; Del Bo *et al.*, 2011 ; Le Ber *et al.*, 2013 ; Hirano *et al.*, 2013 ; Teyssou *et al.*, 2013 ; Pottier *et al.*, 2015 ; Le Ber *et al.*, 2015; Kovacs *et al.*, 2016). Enfin, des mutations dans le gène *TBK1*, qui régule la phosphorylation de p62/SQSTM1 et de l'optineurine, ont récemment été montrées comme responsables de cas de SLA/DFT (Cirulli *et al.*, 2015 ; Freischmidt *et al.*, 2015).

En conclusion, l'autophagie, qui est un mécanisme essentiel à la survie des neurones (Hara *et al.*, 2006 ; Komatsu *et al.*, 2006), semble impliquée à divers niveaux dans la pathogenèse de la SLA/DFT. Un rôle de C9ORF72 dans l'autophagie (cf. partie Résultats ci- dessous), notamment en lien avec TBK1, p62/SQSTM1 et l'optineurine pourrait donc permettre de mieux comprendre ces maladies mais aussi aider à définir des approches thérapeutiques.

III-2- Résultats (revues 3 & 4, article 6)

Une expansion de répétitions G_4C_2 dans le gène *C9ORF72* est la cause la plus fréquente de SLA/DFT. Ces expansions conduisent à une diminution d'expression de la protéine C9ORF72. Toutefois, les fonctions de cette protéine étant mal connues, je me suis intéressée à mieux la caractériser.

III-2-1- Interactants de la protéine C9ORF72

Afin de mieux comprendre les fonctions de la protéine C9ORF72, nous avons cherché les interactants de celle-ci par une analyse protéomique. Après expression dans des cellules Neuro2A de la protéine C9ORF72 fusionnée à une double étiquette Ha-Flag suivie de deux purifications successives sur colonnes d'affinité, différents interactants ont été identifiés par spectrométrie de masse. J'ai pu ainsi mettre en évidence que la protéine C9ORF72 forme un complexe stable avec les protéines SMCR8 et WDR41. WDR41 est une protéine de 52 kDa de fonction inconnue et contenant six motifs WD40 d'interaction protéine-protéine. SMCR8 est une protéine de 105 kDa qui présente, comme C9ORF72, un domaine DENN. SMCR8 interagit avec la kinase ULK1 régulant l'initiation de l'autophagie (Behrends *et al.*, 2010). Toutefois, la fonction de SMCR8 dans l'autophagie est inconnue. De plus, nous avons également mis en évidence que le complexe C9ORF72-SMCR8-WDR41 interagit faiblement et sans doute indirectement avec les protéines p62/SQSTM1 et optineurine, deux récepteurs de l'autophagie faisant le lien entre la membrane de l'autophagosome et les protéines ubiquitinylées.

III-2-2- Activité GEF du complexe C9ORF72-SMCR8-WDR41

Les protéines C9ORF72 et SMCR8 présentent un domaine DENN caractéristique des facteurs d'échange nucléotidique pour les RAB GTPases (Zhang *et al.*, 2012 ; Levine *et al.*, 2013). Mes résultats montrent une interaction entre le complexe C9ORF72-SMCR8 et plusieurs RAB GTPases, dont RAB8A, RAB39B et une interaction plus faible avec RAB6, RAB12, RAB25, RAB33a et RAB38. La protéine C9ORF72 seule n'interagit que très faiblement avec RAB8A et RAB39B. Inversement, SMCR8 seule interagit avec de nombreuses RAB GTPases comme RAB24, RAB32 ou RAB29. Ces interactions sont perdues lorsque SMCR8 est en complexe avec C9ORF72, suggérant ainsi que C9ORF72 modifie la spécificité d'interaction de SMCR8.

Entre autres fonctions, RAB8A et RAB39B régulent la formation et le transport des vésicules d'autophagie (Pilli *et al.*, 2012; Seto *et al.*, 2013; Sato *et al.*, 2014). De plus, RAB8 est impliquée dans un modèle de DFT chez la Drosophile (West *et al.*, 2015) et des mutations dans RAB39 sont responsables d'un retard mental associé à des troubles autistiques et un parkinsonisme juvénile (Giannandrea *et al.*, 2010).

Ces interactions suggèrent que C9ORF72 et/ou SMCR8 pourraient avoir une activité GEF (guanine nucleotide exchange factors). J'ai pu montrer à l'aide de protéines recombinantes purifiées que le complexe C9ORF72-SMRC8-WDR41 présente une activité GEF *in vitro* sur les protéines RAB8A et RAB39B. En effet, ce complexe incubé en présence de protéine RAB8A ou RAB39B pré-chargée en GDP radioactif stimule la libération de GDP. Il est à noter que l'incubation en présence de C9ORF72 seul n'a pas d'effet, suggérant ainsi que seul SMCR8 possède une activité GEF ou que C9ORF72 ne soit actif qu'une fois associé à SMCR8 et WDR41.

III-2-3- C9ORF72 et autophagie

La protéine C9ORF72 interagissant avec différentes protéines impliquées dans l'autophagie, j'ai donc analysé les conséquences de la perte de C9ORF72 sur la formation des autophagosomes. Dans un premier temps, la transfection d'un plasmide GFP-RFP-LC3B dans des cultures primaires corticales d'embryon de souris nous a permis de montrer que la déplétion de la protéine C9ORF72 abolit l'activation de l'autophagie provoquée par l'inhibition de la kinase mTOR. Ces résultats ont été confirmés par western blotting. La diminution d'expression de C9ORF72 dans des cellules neuronales inhibe légèrement la lipidation de LC3B-I en LC3B-II en condition basale. Au contraire, la perte de C9ORF72 abolit presque totalement la lipidation de LC3B-I en condition stimulée par la torine.

Dans une deuxième approche, nous nous sommes intéressés au récepteur de l'autophagie p62/SQSTM1, dont l'accumulation sous forme d'inclusions cytoplasmiques est l'une des caractéristiques histologiques des patients SLA/DFT avec une expansion G_4C_2 dans *C9ORF72* (Al-Sarraj *et al.*, 2011). La déplétion de *C9ORF72* ou de ses partenaires SMCR8 ou WDR41 dans des cellules neuronales conduit à une accumulation d'agrégats de p62/SQSTM1. Il est intéressant de noter que l'expression d'un mutant de RAB39B lié constitutivement au GTP, et donc toujours actif, permet d'outrepasser la déplétion de *C9ORF72* et de supprimer l'accumulation des agrégats de p62/SQSTM1.

III-2-4- Phosphorylation de SMCR8 par TBK1 et ULK1

Nous avons noté que les kinases ULK1 et TBK1, responsables de l'initiation de l'autophagie, étaient toutes deux présentes dans la liste des interactants putatifs de C9ORF72. Toutefois, mes résultats de co-immunoprécipitation montrent que TBK1 n'interagit probablement pas directement avec C9ORF72 ou SMCR8. Cependant, le complexe C9ORF72-SMCR8 interagit avec les adaptateurs de TBK1 : TANK, SINTBAD, NAP1 et l'optineurine. De plus, j'ai pu mettre en évidence que TBK1 phosphoryle SMCR8 sur la sérine 402 et la thréonine 796. Au contraire, TBK1 ne phosphoryle pas C9ORF72, du moins dans nos conditions testées. De même, nous avons confirmé que SMCR8 interagit avec le complexe kinase ULK1-ATG13-ATG101-FIP200 et que SMCR8 peut être phosphorylé par ULK1 sur les serines 400, 492, 562, 667 et la thréonine 666. Tout comme TBK1, ULK1 ne phosphoryle pas C9ORF72 ou WDR41.

J'ai observé que la déplétion de SMCR8 dans des cellules neuronales entraîne une altération de processus autophagique visible par une accumulation d'agrégats de p62/SQSTM1 dans le cytoplasme. De façon intéressante, la surexpression d'un mutant SMCR8 phosphomimétique pour la thréonine 796 (site de phosphorylation de TBK1) permet de compenser la perte de SMCR8 alors que la surexpression d'un mutant phosphomimétique pour les sérines 400, 492, 562 et thréonine 666 (site de phosphorylation de ULK1) ne permet pas de corriger les défauts d'autophagie et l'accumulation d'agrégats p62/SQSTM1. Ces résultats suggèrent donc que la phosphorylation de SMCR8 par TBK1 est un événement important pour la régulation de l'autophagie.

III-2-5- Déplétion de *C9ORF72* et agrégation protéique

L'une des caractéristiques de la SLA/DFT est la présence d'inclusions cytoplasmiques de protéine TDP-43 anormalement phosphorylée, ubiquitinylée et tronquée (Arai *et al.*, 2006 ; Neumann *et al.*, 2006). Les agrégats de TPD-43 étant éliminés par autophagie (Filimonenko *et al.*, 2007 ; Ju *et al.*, 2009 ; Wang *et al.*, 2012), nous avons étudié les conséquences de la perte de la protéine C9ORF72 sur la protéine TDP-43. J'ai ainsi observé que la déplétion de *C9ORF72* entraîne l'apparition de la forme tronquée de 30 kDa de TDP-43 après 7 jours de traitement avec des lentivirus exprimant un shARN contre *C9ORF72*. Au contraire, nous n'observons pas d'accumulation de TDP-43 après 24 ou 48 heures de déplétion de *C9ORF72* en cellules neuronales.

Nous nous sommes alors intéressés à l'ataxine-2, qui lorsqu'elle présente une taille intermédiaire (27 à 33) de répétitions polyQ, augmente le risque de développer une SLA/DFT (Elden *et al.*, 2010 ; Lattante *et al.*, 2014). En condition contrôle, l'expression de plasmide codant l'ataxine-2 contenant 22 polyQ (taille normale) ou 30 polyQ (taille intermédiaire) montre une localisation diffuse de l'ataxine-2 dans le cytoplasme. La déplétion de *C9ORF72* par siARN entraîne l'agrégation de l'ataxine-2 30 polyQ alors qu'elle n'affecte pas l'ataxine-2 avec 22 polyQ. Il est important de noter que nous avons observé des résultats similaires concernant la viabilité cellulaire. En effet, l'expression de l'ataxine-2 avec 22 ou 30 polyQ ou la déplétion de *C9ORF72* seule n'affecte pas la survie des cellules neuronales. Cependant, la déplétion de *C9ORF72*, associée à l'expression de l'ataxine-2 avec 30 polyQ induit une toxicité cellulaire alors qu'elle n'affecte pas la viabilité cellulaire en présence d'ataxine-2 22 polyQ. Ces résultats ont été confirmés par l'équipe d'Edor Kabashi (ICM, Paris) dans un modèle de zébrafish après diminution d'expression de *C9orf72*.

Ces résultats suggèrent donc que la perte seule de C9ORF72 n'induit pas de toxicité cellulaire majeure. Par contre, la diminution d'expression de *C9ORF72* semble conduire à une autophagie sous-optimale qui se révèle inadaptée à un second stress comme l'expression d'ataxine-2 avec des tailles intermédiaires de polyQ.

III-3- Conclusion & discussion

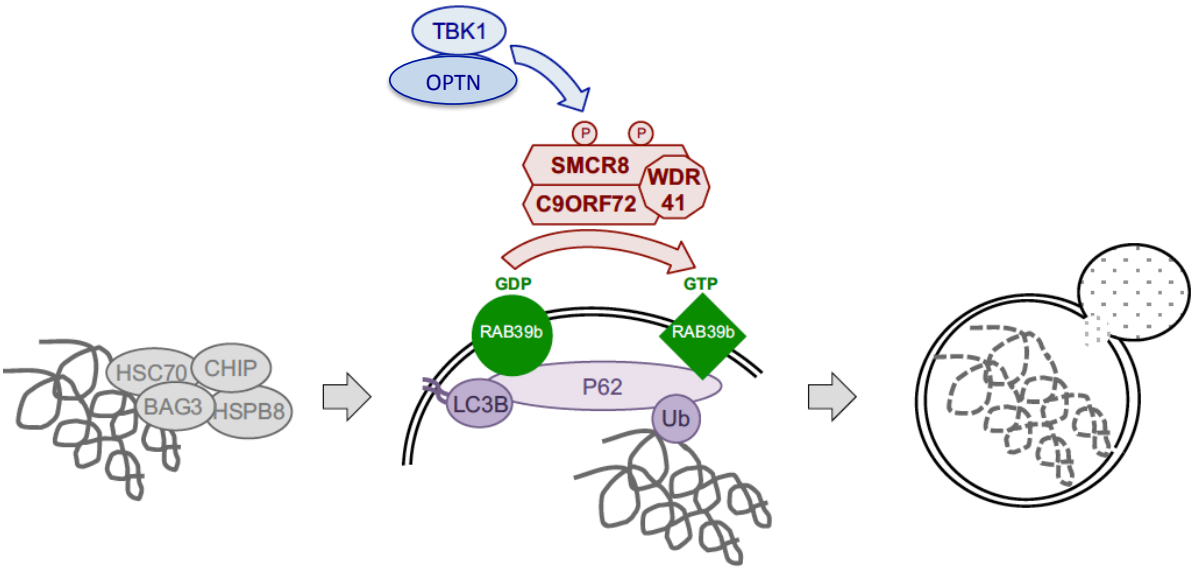
Nos résultats ont montré que la protéine C9ORF72 forme un complexe avec les protéines SMCR8 et WDR41. Ce complexe interagit et régule l'échange GDP/GTP de plusieurs protéines RAB GTPases, dont notamment RAB8A et RAB39B. Enfin, le complexe C9ORF72-SMCR8-WDR41 régule l'autophagie et la perte de C9ORF72 sensibilise les neurones à un second stress comme l'ataxine-2 avec des tailles intermédiaires de polyQ.

Depuis, plusieurs équipes ont mis en évidence l'implication de C9ORF72 dans la régulation de l'autophagie (Webster *et al.*, 2016 ; Sullivan *et al.*, 2016; Yang *et al.*, 2016). De même, l'interaction de C9ORF72 avec SMCR8 et WDR41 a été confirmée par plusieurs études (Sullivan *et al.*, 2016; Amick *et al.*, 2016; Yang *et al.*, 2016 ; Xiao *et al.*, 2016). Il est à noter que si les travaux de Yang et collaborateurs (2016) ont montré, en accord avec nos résultats, que le complexe C9ORF72 interagit avec RAB39B, l'étude de Webster et collaborateurs a identifié une interaction entre C9ORF72 et RAB1 liée au GTP, alors que Farg et collaborateurs ont montré que C9ORF72 interagit avec les RAB GTPases RAB1, RAB5, RAB7 et RAB11 (Farg *et al.*, 2014 ; Webster *et al.*, 2016). Cependant, dans notre étude, nous n'avons pas retrouvé ces résultats, ce qui peut s'expliquer notamment par le fait que les types cellulaires et le protocole d'immunoprécipitation que nous avons utilisé sont différents. De plus, nous avons essentiellement étudié l'interaction du complexe C9ORF72-SMCR8 et n'avons pas étudié RAB1 lié au GTP ou la protéine C9ORF72 seule. Enfin, il existe plus d'une soixantaine de RAB GTPases et il est hautement probable que le complexe C9ORF72-SMCR8-WDR41 régule différentes protéines RAB-GTPases selon le type cellulaire et/ou les conditions cellulaires.

L'interaction du complexe C9ORF72 avec le complexe ULK1 a été confirmée par plusieurs études, toutefois sans préciser le rôle ou l'importance de cette interaction (Behrends *et al.*, 2010 ; Webster *et al.*, 2016; Sullivan *et al.*, 2016; Yang *et al.*, 2016). Nous avons montré que SMCR8 est phosphorylée par ULK1 et TBK1 sur des résidus différents. Des mutants phosphomimétiques mimant une phosphorylation constitutive de SMCR8 par TBK1 corrige l'altération d'autophagie provoquée par la perte de SMCR8 alors que des mutants phosphomimétiques mimant la phosphorylation de SMCR8 par ULK1 sont sans effet. De plus, un mutant constitutivement actif de RAB39B permet de corriger les altérations d'autophagie provoquées par la déplétion de *C9ORF72*, de SMCR8 ou même de TBK1. Ces résultats suggèrent, d'une part, que TBK1 joue un rôle essentiel dans la régulation du processus d'autophagie dans les

neurones, et d'autre part, que les protéines TBK1, C9ORF72-SMCR8-WDR41 et RAB39B appartiennent à une même voie de régulation de l'autophagie (figure 14). Nous proposons donc un modèle où la kinase TBK1 phosphoryle SMCR8 en complexe avec C9ORF72 et WDR41. Ce complexe agit comme une GEF pour les RAB8A et RAB39B, lesquelles pourraient être impliquées dans les étapes précoces de la biogenèse de l'autophagosome. Cependant, il reste à déterminer comment la kinase TBK1 est régulée, ainsi que le rôle précis de la phosphorylation de SMCR8 par TBK1. De la même manière, il serait intéressant de préciser le rôle des RAB8A et RAB39B dans le processus d'autophagie.

***Figure 14** : Schéma montrant l'implication possible du complexe C9ORF72-SMCR8 dans le processus autophagique.*



Protéines mal conformées ou agrégées (TDP43, ATX2…)

Formation de l'autophagosome

Fusion de l'autophagosome et du lysosome

L'une des caractéristiques de la grande majorité des patients sporadiques atteints de SLA/DFT est la présence d'inclusions contenant la protéine TDP-43 phosphorylée, ubiquitinylée et tronquée (Neumann *et al.*, 2006). Ces agrégats sont également détectés dans les formes familiales avec des mutations dans le gène *TARDBP* ou dans d'autres gènes impliqués dans l'autophagie comme *TBK1*, *SQSTM1*, *UBQLN2*, *OPTN*, *C9ORF72* et *GRN* (Marangi and Traynor,

2015). Nous avons observé que la déplétion de *C9ORF72* conduit à une autophagie sous-optimale associée à l'apparition de la forme tronquée et d'agrégats de TDP-43 dans des cultures primaires corticales de neurones. Ces résultats sont consistants avec plusieurs études montrant qu'une altération de l'autophagie conduit à l'accumulation d'agrégats de TDP-43 (Filimonenko *et al.*, 2007 ; Ju *et al.*, 2009; Urushitani *et al.*, 2010; Wang *et al.*, 2012; Barmada *et al.*, 2014; Scotter *et al.*, 2014). Il est donc tentant de proposer un lien causal direct entre les mutations de ces gènes impliqués dans l'autophagie et la présence d'inclusions de TDP-43. Cette hypothèse soulève alors la question de la cause des agrégats de TDP-43 dans les SLA sporadiques, et notamment si l'efficacité de l'autophagie pourrait être diminuée chez ces patients.

Enfin, nous avons observé que la seule perte de *C9ORF72* n'est pas ou peu toxique. Au contraire, la diminution d'expression de *C9ORF72* associée à une taille intermédiaire de répétitions polyQ dans l'ataxine-2 conduit à une mort neuronale. Ces résultats sont intéressants au regard des modèles de souris knockout pour C9ORF72 qui développent des dysfonctions immunitaires mais ne montrent pas de neurodégénérescence (Koppers *et al.*, 2015 ; O'Rourke *et al.*, 2016; Atanasio *et al.*, 2016 ; Sudria-Lopez *et al.*, 2016; Burberry *et al.*, 2016; Jiang *et al.*, 2016). Il pourrait donc être intéressant de tester dans ces souris si la perte de *C9ORF72* n'est pas associée à un phénotype neurodégénératif uniquement en présence d'un second stress comme celui provoqué par des tailles intermédiaires de polyQ dans l'ataxine-2 ou l'expression de DPR ou d'ARN G_4C_2. Dans ce modèle, la diminution d'expression de *C9ORF72* ne serait donc pas un événement pathogène primaire mais sensibiliserait les neurones à un deuxième stress apparaissant avec l'âge. Cette hypothèse serait cohérente avec l'absence de mutations « classiques » perte de fonction (codon stop prématuré, mutation des sites d'épissages, délétions, etc.) dans le gène *C9ORF72*.

III-4- Projet de recherche

III-4-1- Rôle de C9ORF72 et SMCR8 dans l'autophagie en lien avec la SLA/DFT

L'autophagie est essentielle pour les neurones et sa suppression dans la souris entraîne une neurodégénération (Hara *et al.,* 2006). De plus, des inclusions cytoplasmiques contenant LC3-II, OPTN et p62/SQSTM1, des marqueurs caractéristiques des autophagosomes, ont été détectées dans les motoneurones de patients SLA sporadiques (Sasaki *et al.*, 2011). Enfin, plusieurs gènes (*TBK1*, *UBQLN2*, *CHMP2B*, *VCP*, *p62/SQSTM1*, *OPTN,* etc*.)* mutés dans la SLA/DFT sont impliqués dans l'autophagie ou la dégradation des protéines. L'implication de C9ORF72 dans la régulation de l'autophagie pourrait donc être importante pour mieux comprendre la SLA/DFT. Je souhaite donc explorer davantage les liens entre le complexe C9ORF72-SMCR8-WDR41 et la régulation de l'autophagie.

Nous avons développé des anticorps monoclonaux contre C9ORF72 et SMCR8 et mes résultats les plus récents montrent que ces protéines forment de petites vésicules dans le cytoplasme des cellules neuronales en culture. De plus, j'observe une co-localisation partielle (20%) de ces vésicules avec LAMP2, un marqueur des lysosomes. La déplétion de *C9ORF72* par siARN entraîne une perte de localisation de SMCR8 aux lysosomes, confirmant que C9ORF72 est important pour la stabilité du complexe C9ORF72-SMCR8 mais peut-être aussi pour l'ancrage de ce complexe aux lysosomes. Dans un premier temps, je souhaite poursuivre ces recherches et mieux comprendre à quelle étape de l'autophagie le complexe C9ORF72-SMCR8 agit. Pour cela, je propose d'étudier la co-localisation de C9ORF72 avec les RAB-GTPases RAB8A et RAB39B et avec divers marqueurs des autophagosomes comme LC3B-II, WIPI2, ATG16L,etc.) par microscopie confocale mais aussi par microscopie super résolutive (Leica SR GSD). Je propose aussi d'étudier si la perte d'expression par siARN de *C9ORF72* ou *SMCR8* perturbe la formation, le transport ou la fusion des autophagosomes aux lysosomes dans des cultures neuronales. Enfin, la fonction de RAB8A et RAB39B est mal connue dans l'autophagie. J'aimerais donc aussi tester l'effet de siARN contre RAB8A et RAB39B sur la formation, le transport et la fusion des autophagosomes, mais aussi chercher les interactants de RAB8A et RAB39B par la technique de purification en tandem HA-Flag.

La protéine C9ORF72 forme un complexe avec les protéines SMCR8 et WDR41 (Sullivan *et al.*, 2016; Amick *et al.*, 2016; Yang *et al.*, 2016; Xiao *et al.*, 2016). Ce complexe C9ORF72-SMCR8-WDR41 régule l'autophagie et agit comme facteur d'échange GDP/GTP pour les RAB GTPases RAB8A et RAB39B. Toutefois, il est probable que les protéines C9ORF72 et SMCR8 agissent aussi seules ou dans d'autres complexes. En effet, nous avons développé des souris knockout pour *C9orf72* ou *Smcr8* et ces souris ne présentent pas le même phénotype. Ainsi, le knockout du gène *C9orf72* conduit à des problèmes immunitaires caractérisés par une adénopathie, une splénomégalie, une forte inflammation et une hyperproduction d'anticorps, phénotype qui a été également observé par d'autres équipes (O'Rourke *et al.*, 2016 ; Atanasio *et al.*, 2016; Sudria-Lopez *et al.*, 2016 ; Burberry *et al.*, 2016; Jiang *et al.*, 2016). Au contraire, les souris knockout pour le gène *Smcr8* présentent une petite taille sans signe évident d'inflammation (du moins à 5 mois d'âge). Nous allons explorer plus en détail le phénotype de ces souris *Smcr8*$^{-/-}$, mais ces premiers résultats suggèrent que C9ORF72 et SMCR8 puissent avoir des fonctions distinctes. Plusieurs protéines telles que NPRL2 et NPRL3 sont structurellement proches de C9ORF72 (Zhang *et al*., 2012). Par des expériences de co-immunoprécipitation, j'ai observé que SMCR8 interagit avec NPRL2 ou NPRL3 (environ 50 à 70 % de l'interaction avec C9ORF72) et que l'expression de NPRL3 dans des neurones en culture permet de supprimer l'accumulation d'agrégats de p62/SQSTM2 entrainée par la perte par siARN contre *C9ORF72*. Ces résultats suggèrent que SMCR8 puisse former un complexe actif soit avec C9ORF72, soit avec NPRL3. J'aimerais poursuivre l'étude de ce complexe, notamment par la recherche d'interactants du complexe SMCR8-NPRL3 en utilisant la même approche (double étiquette FLAG-HA) que celle développée pour C9ORF72. De plus, NPRL3 est localisé au lysosome et régule la voie mTOR (Bar-Peled *et al.*, 2013) dont une altération conduit à de nombreuses altérations cellulaires et dont la diminution d'expression chez la souris conduit à des animaux de petite taille (Zhang *et al.*, 2011). Il serait donc intéressant de tester un rôle potentiel de SMCR8 dans la voie mTOR, et notamment de relier cette voie de signalisation au phénotype des souris knockout pour *Smcr8*. De plus, SMCR8 est identifié parmi les protéines phosphorylées par le complexe mTOR1 (Hsu *et al.,* 2011), il serait intéressant d'étudier les conséquences de cette phosphorylation. Je serai aidée dans ce travail par Camille Corbier, une étudiante qui commence sa thèse et que je co-encadre. Enfin, ce travail pourrait suggérer que la fonction de C9ORF72 puisse être partiellement compensée par NPRL3 dans le cerveau de souris knockout pour *C9orf72*, expliquant ainsi l'absence de phénotype

neuronal dans ces souris. Il serait donc intéressant de croiser des souris knockout pour *C9orf72* avec des souris hétérozygotes pour *Nprl3* ou avec une perte de ce gène uniquement dans le cerveau (le knockout total de *Nprl3* étant létal ; Kowalczyk *et al.*, 2012).

En parallèle, l'équipe du Professeur Jan Veldink (UMC Utrecht Brain Center, Pays-Bas) a identifié plusieurs mutations de *SMCR8* dans une cohorte de patients SLA/DFT. L'analyse préliminaire des conséquences de ces mutations montre que celles-ci agissent comme des dominants négatifs de l'activité du complexe C9ORF72-SMCR8. Je propose donc d'étudier l'effet de ces mutations dans le gène *SMCR8* sur la fonction du complexe C9ORF72-SMCR8 et sur l'autophagie (purification en tandem HA-Flag des protéines associées aux mutants de *SMCR8*, effet des mutants sur la formation, transport et fusion des autophagosomes par microscopie, etc.). En cas de confirmation, il s'agirait d'un résultat important car la découverte de mutations, même rares, de SMCR8 dans la SLA/DFT renforcerait fortement l'hypothèse d'une altération de l'activité de ce complexe et de l'autophagie dans ces maladies.

III-4-2- Rôle de l'ataxine-2 dans la SLA/DFT en lien avec C9ORF72 :

La présence de taille intermédiaire de polyglutamines dans l'ataxine-2 est un modificateur fort de la SLA/DFT (Elden *et al.,* 2010 ; Lattente *et al.,* 2014). Cependant, il reste à comprendre comment et par quels mécanismes ces expansions de polyQ agissent dans la SLA/DFT. Des études montrent que l'ataxine-2 régule la stabilité des ARN messagers en interagissant avec la protéine PABP et l'ARN *via* un domaine Lsm/LsmAD (Satterfield and Pallanck, 2006). Toutefois, le lien entre cette fonction de stabilisation des ARN messagers et la SLA/DFT n'est pas compris. De plus, mes résultats montrent que la diminution d'expression de *C9ORF72* dans des cellules neuronales provoque l'agrégation et la toxicité de l'ataxine-2 contenant 30 polyQ (taille intermédiaire). Au contraire, l'ataxine-2 avec une taille de polyQ normale (22 Q) reste diffuse et non toxique (figure 15, Sellier *et al.,* 2016).

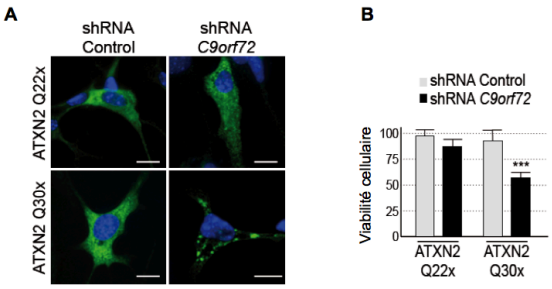


Figure 15 : *Effet de la déplétion de la protéine C9ORF72 sur l'ataxine-2 avec une taille normale ou intermédiaire de polyQ. (A) Localisation de l'ataxine-2 transfectée dans des cellules neuronales. (B) Analyse de la toxicité provoquée par l'ataxine-2 avec une taille normale ou intermédiaire de polyQ après déplétion ou non de C9ORF72 dans des cellules neuronales.*

Mes résultats suggèrent donc un lien entre les protéines C9ORF72 et l'ataxine-2. J'aimerais poursuivre ce travail et notamment étudier comment l'ataxine-2 avec des tailles intermédiaires de polyQ peut être toxique pour les neurones. Pour cela, je souhaite identifier par UV-CLIP séquençage les ARNm neuronaux liés par l'ataxine-2 avec une taille normale ou intermédiaire de polyglutamines. Je souhaite aussi identifier par purification en tandem HA-Flag et analyse protéomique les interactants de l'ataxine-2 avec une taille normale ou intermédiaire de polyQ. De plus, en construisant différents mutants de l'ataxine-2, je souhaite mieux comprendre par quel mécanisme la diminution de la protéine C9ORF72 entraîne une agrégation spécifique de l'ataxine-2 alors qu'elle n'affecte pas la localisation ou l'agrégation d'autres protéines contenant des expansions de polyQ comme la Huntingtine ou l'Ataxine-3. Enfin, le Professeur Auburger (Goethe University Medical School, Francfort, Allemagne) a développé une souris knockin présentant 42 glutamines dans le gène *Atxn2*. Tout comme les souris knockout *C9Orf72*, ces souris ne présentent pas de phénotypes neurodégénératifs évidents (Damrath *et al.*, 2012). Afin d'étudier une synergie entre la perte de C9ORF72 et l'expression de polyQ intermédiaires dans l'ataxine-2, je propose de croiser des souris présentant un knockout conditionnel pour le gène *C9Orf72* avec des souris exprimant la recombinase CRE sous la dépendance d'un promoteur Nestine pour obtenir une perte de *C9orf72* spécifiquement dans le cerveau. Ces souris seront alors croisées avec les souris ataxine-2 42 polyQ et leur descendance sera testée pour identifier d'éventuelles altérations de leurs fonctions motrices et/ou cognitives, des signes histopathologiques et moléculaires de neurodégénérescence, etc. Cette approche devrait permettre de mieux comprendre si l'expression de l'Ataxine-2 avec une taille intermédiaire de polyQ modifie, et notamment augmente, la toxicité due à la perte d'expression de *C9Orf72*.

III-4-3- Recherche de molécules pharmacologiques

J'ai observé que la diminution d'expression d'environ 50% de *C9ORF72* par shARN dans des neurones en culture conduit à l'accumulation d'agrégats de p62/SQSTM1, de LC3B et d'ataxine-2 contenant 30 polyQ, suggérant une altération de l'autophagie dans ces neurones.

Ces agrégats de p62/SQSTM1, qui constituent un marqueur pathologique des patients avec une expansion G_4C_2 dans *C9ORF72* (Al-Sarraj *et al.*, 2011), sont facilement observables et quantifiables. De plus, l'autophagie est un mécanisme cellulaire hautement régulé et qui peut être modulé par divers traitements pharmaceutiques (Rubinsztein *et al.*, 2007). Je souhaite donc développer un criblage de composés pharmacologiques automatisé afin d'identifier des molécules corrigeant les altérations de l'autophagie dues à une perte de C9ORF72. Pour cela, j'utiliserai des cellules neuronales immortalisées GT17, mieux adaptées que les cultures primaires à un criblage à haut débit. J'ai confirmé que la perte de C9ORF72 conduit aussi à des agrégats de p62/SQSTM1 dans ces cellules. Je souhaite donc maintenant automatiser les différentes étapes de ce test en collaboration avec la plateforme de criblage à haut débit de l'IGBMC. Nous criblerons dans un premiers temps la banque Prestwick, constituée d'environ 1200 molécules utilisées en clinique humaine, puis, selon les résultats, nous testerons la banque de composés de la chimiothèque de Strasbourg (6000 molécules – laboratoire PCBIS) et éventuellement la chimiothèque nationale (jusqu'à 65000 composés et substances naturelles). Selon les résultats, je testerais les molécules les plus prometteuses sur des neurones différenciés à partir de cellules iPS de patients SLA/DFT (collaboration avec le Prof Fen-Biao Gao, University of Massachusetts, USA et avec le Dr Martinat Cécile, i-Stem, Evry) puis éventuellement sur les modèles souris (C9ORF72 knockout / ataxine-2 polyQ) que je développe.

C9ORF72 is a GDP/GTP exchange factor for Rab8 and Rab39 and regulates autophagy.

Camille Corbier[1], Chantal Sellier[1#]

[1]Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), INSERM U964, CNRS UMR7104, Strasbourg University, Illkirch, France.

Correspondance to Chantal.sellier@igbmc.fr; +33 (0)3 88 65 33 57

ABSTRACT

Amyotrophic Lateral Sclerosis and Frontotemporal Dementia (ALS-FTD) are devastating neurodegenerative disease affecting motoneurons from the spinal chord and neurons from the frontal and temporal cortex, respectively. The most common genetic cause for ALS-FTD is an expansion of GGGGCC repeats within the first intron of the C9ORF72 gene. However, little is known on the function of C9ORF72. Recently, other and we found that C9ORF72 forms a stable complex with the SMCR8 and WDR41 proteins. This complex acts as a GDP/GTP exchange factor for the small RAB GTPases Rab8a and Rab39b. Since Rab8 and Rab39 are involved in macroautophagy, we tested the role of C9ORF72 in this mechanism. Decrease expression of C9ORF72 in neuronal cultures leads to autophagy dysfunction characterized by accumulation of aggregates of p62/SQSTM1. However, loss of C9ORF72 expression does not cause major neuronal cell death, suggesting that a second stress may be required to promote cell toxicity. Intermediate size of polyglutamine repeats within Ataxin-2 (ATXN2) is an important genetic modifier of ALS-FTD. We found that decrease expression of C9ORF72 synergizes the toxicity and aggregation of ATXN2 with intermediate size of polyglutamine (30Q). Overall, our data suggest that reduce expression of C9ORF72 causes suboptimal autophagy that sensitizes neurons to a second stress. These data suggest that reduce expression of C9ORF72 may partly contribute to ALS-FTD pathogenesis.

Keywords

Amytrophic Lateral Sclerosis (ALS); C9ORF72; Rab8; Rab39

Introduction

Amyotrophic lateral sclerosis (ALS) is a fatal motoneuronal neurodegenerative disorder for which no effective treatment is available. With an estimated incidence of 2.6/100,000 per year and a lifetime risk of 1:350 for male and 1:450 for female, Amyotrophic lateral sclerosis (ALS) is the third most common neurodegenerative disease worldwide and the most frequent motor neuron disease. ALS is mainly characterized by the massive degeneration of both upper (UMN) and lower (LMN) motor neurons in the cerebral cortex and spinal cord, which rapidly leads in most patients to paralysis and death due to denervation of the respiratory muscles. Frontotemporal dementia (FTD) is the second most common presenile dementia after Alzheimer's disease. Increasing evidences indicate that FTD and ALS pathologies form a continuum of neurological diseases, which share a common pathological background. First, ALS and FTD patients have an overlap of clinical symptoms, since approximately 15% of FTD patients have motor dysfunction meeting the criteria of ALS and 15% to 30% of ALS patients have FTD. Also, histopathological analyses demonstrate that ALS and FTD patients share common histopathological markers with accumulation of cytoplasmic aggregates of phosphorylated and cleaved transactive response DNA-binding protein 43 (TDP-43) in the vast majority of ALS patients and in the most common Tau-negative pathological subtype of FTD (Neumann et al., 2006). Finally, the concept that FTD and ALS represent a clinicopathological spectrum of disease is confirmed by genetic evidences as mutations in *TARDBP* (encoding TDP-43 protein), *UBQLN2*, *FUS* and, topic of this highlight, *C9ORF72*, lead to co-occurrence of ALS-FTD. Importantly, an expansion of hundreds to thousands of GGGGCC repeats within the first intron of the C9ORF72 gene represents the most common inherited cause for ALS and FTD, accounting for 20-60% of familial forms and 1-7% of sporadic ALS-FTD patients in Northern Europe and North America (DeJesus-Hernandez et al., 2011; Renton et al., 2011). The GGGGCC expansion is located between two five prime non-coding exons of C9ORF72, which encodes a poorly characterized protein. Three main non-exclusive mechanistic models of C9ORF72-mediated ALS have been proposed (review in Haeusler et al., 2016).

• First, various studies indicate that expanded GGGGCC repeats are transcribed both in the sense and the antisense strands, forming nuclear aggregates of RNA containing expanded GGGGCC or GGCCCC repeats. Such mutant RNA may bind and titrate specific RNA binding proteins, including hnRNP-A3, Pur-α, ADARB2, hnRNP-H, and Nucleolin (Almeida et al., 2013; Donnelly et al., 2013; Haeusler et al., 2014). However, whether the recruitment of these RNA binding proteins leads to titration and loss of their function remains to be determined.

• The second potential mechanism for neurotoxicity of expansions is a form of non-canonical protein translation termed repeat-associated, non-ATG (RAN) translation (Zu et al., 2011). Extensive studies have now established that sense and antisense RNA containing the expanded repeats are RAN translated in all six sense and antisense frames, resulting in expression of five different di-peptide repeats containing proteins (DPRs), which form inclusions throughout the brain of patients with C9-ALS/FTD (Ash et al., 2013; Gendron et al., 2013; Mori et al., 2013; Zu et al., 2013). These DPRs were also identified in mice expressing expanded GGGGCC repeats (Chew et al., 2015; Peters et al., 2015; O'Rourke et al., 2015; Liu et al., 2016; Jiang et al., 2016), and were found to be toxic in neuronal cell cultures and in Drosophila

models through alteration of the nucleocytoplasmic transport (Tao et al., 2015; Freibaum et al., 2015; Zhang et al., 2015; Jovičić et al., 2015).

• Third, several studies consistently found reduced levels of C9ORF72 transcripts in GGGGCC expanded-repeats carriers, suggesting a possible loss-of-function disease mechanism (DeJesus-Hernandez et al., 2011; Renton et al., 2011; Almeida et al., 2013; Waite et al., 2014; van Blitterswijk et al., 2015). Haploinsufficiency of C9ORF72 is to be anticipated if the hexanucleotide-expanded genomic DNA promotes epigenetics modification and impairs C9ORF72 transcription. However, the absence of neuronal phenotypes in mouse depleted of C9orf72 expression in brain or in neurons (Lagier-Tourenne et al., 2013; Koppers et al., 2015), as well as the absence of ALS/FTD patients with null alleles or misense mutations in C9ORF72, argue against a loss-of-function of C9ORF72 as the sole or main cause of ALS-FTD. Nevertheless, evidences of locomotion deficit in zebrafish with reduced expression of C9ORF72 (Ciura et al., 2013), and the correlation of decreased C9ORF72 mRNA expression with decreased patient survival (Van Blitterswijk et al., 2015) suggest that reduced expression of C9ORF72 may partly contribute to ALS-FTD pathogenesis.

C9ORF72 in complex with SMCR8 regulates autophagy through Rab8 and Rab39

Despite many major recent advances, little is known on the normal molecular and cellular functions of C9ORF72. Other and we found that C9ORF72 forms a complex with two proteins, SMCR8 and WDR41 of unknown functions (Sellier et al., 2016; Sullivan et al., 2016; Xiao et al., 2016). Bioinformatics analysis identified that both C9ORF72 and SMCR8 contain DENN (Differentially Expressed in Normal and Neoplastic cells) domains characteristic of Rab GDP/GTP exchange factors (GEFs) (Levine et al., 2013; Zhang et al., 2012). Consistent with these predictions, we found that SMCR8 interacts with various RAB GTPases and that the C9ORF72-SMCR8 complex promotes *in vitro* GDP/GTP exchange for the Rab8a and Rab39b. Since Rab8 and Rab39 are involved in macroautophagy (Pilli et al., 2012; Seto et al., 2013; Sato et al., 2014) and as SMCR8 was identified in proteomic analysis of autophagy network (Behrends et al., 2010), we investigated whether C9ORF72 was regulating autophagy. Depletion of C9ORF72 expression by shRNA and/or siRNA in transformed neuronal cells or in primary cultures of cortical neurons of E18 mouse embryo has a deleterious role on autophagy, with notable accumulation of unresolved aggregates of the autophagy receptor P62/SQSTM1. A function of C9ORF72 in autophagy is consistent with a previous report of LC3B alteration in C9ORF72 siRNA-depleted neuronal cells (Farg et al., 2014), but also with the increased accumulation of P62 and the susceptibility to autophagy inhibitors observed in cultures of human neurons derived from iPS cells of ALS-FTD patients carrier of an GGGGCC expansion in C9ORF72 (Almeida et al., 2013; Dafinca et al., 2016). Furthermore, the recent demonstration that siRNA-mediated depletion of C9ORF72 leads to a decrease formation of LC3B-positive vesicles conclusively pinpoints a role of C9ORF72 in the initiation of autophagy (Webster et al., 2016). Finally, a role of the C9ORF72-SMCR8 complex in autophagy is also strengthened by the observation that this complex interacts with two kinases regulating autophagy, ULK1 and TBK1 (Behrends et al., 2010; Sellier et al., 2016; Sullivan et al., 2016; Webster et al., 2016). Our in vitro phosphorylation assays indicate that both ULK1 and TBK1 kinases phosphorylate SMCR8, but not C9ORF72. Interestingly, a mutant of SMCR8 threonine 796 in aspartic acid,

which mimics a constitutive phosphorylation of SMCR8 by TBK1, is able to correct autophagy dysfunctions caused by siRNA-mediated reduced expression of either SMCR8 or TBK1. This may be relevant to disease pathogenesis as loss of function mutations in TBK1 cause ALS-FTD (Cirulli et al., 2015; Freischmidt et al., 2015). Furthermore, TBK1 is known to phosphorylate P62 and OPTN, two autophagy receptors also found muted in rare case of ALS (Maruyama et al., 2010). In that aspect, we found that Rab39b interacts with P62 and confirmed previous observations (Hattula and Peränen, 2000; Pilli et al., 2012) that OPTN interacts with Rab8a. These results support a model where P62 or OPTN autophagy receptors act as essential hubs to gather specific Rab GTPases with there specific GEF effectors and kinase regulators to initiate autophagy precisely at the site of ubiquitinated protein aggregates, dysfunctional organelles or intracellular pathogens (Figure 1). A model supported by the recent report of the importance of TBK1 recruitment to dysfunctional mitochondria to phosphorylate P62 or OPTN and initiate mitophagy (Matsumoto et al., 2015; Heo et al., 2015; Lazarou et al., 2015). Finally, we found that a mutant form of Rab39b, which is locked in its GTP conformation and does not consequently require any GEF activity, can rescue autophagy dysfunction caused by siRNA-mediated loss of C9ORF72-SMCR8 or of its upstream kinase regulator, TBK1. These results suggest that TBK1, C9ORF72-SMCR8 complex and RAB39b belong to a common pathway regulating autophagy in neuronal cells.

Loss of C9ORF72 synergizes ATXN2 polyQ toxicity

While reduce expression of C9ORF72 causes a partial dysfunction of autophagy, this was not sufficient to trigger overt neuronal cell death in neuronal cultures. This is consistent with the absence of neurodegeneration in mouse depleted of C9orf72 expression in brain or in neurons (Lagier-Tourenne et al., 2013; Koppers et al., 2015). Thus, we hypothesized that a second stress might be required to trigger neuronal cell loss upon C9ORF72 depletion. Intermediate size of 27 to 33 glutamines in Ataxin-2 increases the risk of ALS-FTD (Elden et al., 2010). Importantly, we found that decrease expression of C9ORF72 synergizes the toxicity of Ataxin-2 with 30 glutamines both in mammalian neuronal cell cultures and in zebrafish embryos. Importantly, this synergic or double hit model of toxicity is consistent with the absence of ALS/FTD patients with null or misense mutations in C9ORF72, while there is increasing genetic evidences of oligogenicity in ALS-FTD (Ferrari et al., 2012; Van Blitterswijk et al., 2012). Also, this synergic toxicity appears specific to Ataxin-2 since reduced expression of C9ORF72 does not accentuate aggregation of Ataxin-3 or huntingtin with expanded polyglutamine repeats (Sellier et al., 2016). This is rather unexpected, as these polyglutamine-containing proteins are cleared by autophagy. However, it is possible that the little effect of C9ORF72 loss on the aggregation of Ataxin-3 or huntingtin that we observed is due to the inherent limitation of *in vitro* cell cultures and the short time frame of our study. Alternatively, the aggregations of Htt or Ataxin-3 with expanded polyglutamine repeats may have reach a maximum in our cell culture and cannot be enhanced further. Thus, an effect of C9ORF72 depletion on polyglutamine-containing proteins remains to be tested in animal model and/or on a longer time period of analysis.

Conclusion and perspectives

Expansion of GGGGCC repeats within the first intron of the C9ORF72 gene is the prime cause of ALS-FTD, but little is known on the molecular function of C9ORF72. Recent evidences indicate that C9ORF72 interacts with SMCR8 and

that this complex regulates autophagy (Behrends et al., 2010; Farg et al., 2014; Sellier et al., 2016; Sullivan et al., 2016; Xiao et al., 2016; Webster et al., 2016). However, the precise role of C9ORF72 or SMCR8 in the autophagy pathway remains to be determined. C9ORF72 and SMCR8 in complex or in isolation interact with various Rab GTPases, including Rab8 and Rab39 as well as Rab1, Rab5, Rab7 and Rab11 (Farg et al., 2014; Webster et al., 2016), but it is unclear which Rab would regulate which autophagy steps in which tissue or cellular type. Furthermore, Rab GTPases act in cascades where the passage from one Rab to the next requires the recruitment of Rab GAP and GEF effectors to the upstream and downstream Rab GTPases, respectively. Thus, it remains to determine the importance of C9ORF72-SMCR8 in such Rab cascade as well as the identity of such Rab GTPases. In that aspect, the recent demonstration that Rab1 may be the initial Rab GTPase recruiting C9ORF72 to promote GDP/GTP exchange and thus activation of a downstream Rab GTPase is especially exciting (Webster et al., 2016). It is also highly possible that C9ORF72 and SMCR8 in complex or in isolation bind specific Rab GTPases involved in other cellular process than autophagy. In that aspect, functions of C9ORF72 in endocytosis and lysosomal pathways have been recently suggested (Farg et al., 2014; Sullivan et al., 2016; Busch et al., 2016; O'Rourke et al., 2016). Next, we found that SMCR8 is phosphorylated by ULK1 and TBK1. Also, SMCR8 is found phosphorylated by AMPK and mTOR kinases in large proteomic screen. However, the physiological consequences of SMCR8 phosphorylation by ULK1, TBK1, AMPK or mTOR kinases remain to be explored. Finally, reduce expression of C9ORF72 is not overly toxic by itself but synergizes the toxicity of Ataxin-2 with intermediate size of polyglutamine. This synergic model supports a two hit hypothesis in ALS-FTD and partly explains why the sole loss of C9ORF72 is not sufficient to cause massive neuronal cell death in cell culture or in knockout mouse models. However, it remains to test validity of this model in mammals, notably by testing whether expression of Ataxin-2 with intermediate size of polyglutamine promotes neuronal cell death in C9ORF72 knockout mice. Similarly, it remains to explore whether reduced expression of C9ORF72 may synergize other cellular stress such as RAN translation of expanded GGGGCC repeats into toxic DPRs. Furthermore, as Rab39b interacts with P62 that is involved in the autophagic clearance of protein aggregates, while Rab8 interacts with OPTN that is involved in the autophagy of dysfunctional mitochondria (Matsumoto et al., 2015; Heo et al., 2015; Lazarou et al., 2015), it will be exciting to test whether C9ORF72 is also involved in mitophagy. Last but not least, it is striking to note that various mutations causing ALS-FTD are found in genes involved in protein clearance pathways, including *UBQLN2*, *CHMP2B, VCP, OPTN, SQSTM1* and *TBK1*. Thus, our work linking decrease expression of C9ORF72 to suboptimal autophagy may provide further support to compromised protein-clearance mechanisms in ALS-FTD (Figure 2). Whether suboptimal protein degradation may contribute to disease pathogenesis in ALS-FTD is an exciting area that remains to be fully explored.

ACKNOWLEDGMENTS

This work was supported by Fondation de France Thierry Latran #57486 “Model-ALS”, AFM grant #18605 “Role of C9ORF72 in ALS-FTD”, ERC-2012-StG #310659 “RNA DISEASES”, ANR-10-LABX-0030-INRT and ANR-10-IDEX-0002-02.

REFERENCES

Almeida S, Gascon E, Tran H, Chou HJ, Gendron TF, Degroot S, Tapper AR, Sellier C, Charlet-Berguerand N, Karydas A, Seeley WW, Boxer AL, Petrucelli L, Miller BL, Gao FB. Modeling key pathological features of frontotemporal dementia with C9ORF72 repeat expansion in iPSC-derived human neurons. Acta Neuropathol. 2013 Sep;126(3):385-99.

Ash PE, Bieniek KF, Gendron TF, Caulfield T, Lin WL, van Blitterswijk MM, Jansen-West K, Paul JW 3rd, Rademakers R, Boylan KB, Dickson DW, Petrucelli L. Unconventional translation of C9ORF72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. Neuron. 2013 Feb 20;77(4):639-46.

Behrends C, Sowa ME, Gygi SP, Harper JW. Network organization of the human autophagy system. Nature. 2010 Jul 1;466(7302):68-76.

Chew J, Gendron TF, Prudencio M, Sasaguri H, Zhang YJ, Castanedes-Casey M, Lee CW, Jansen-West K, Kurti A, Murray ME, Bieniek KF, Bauer PO, Whitelaw EC, Rousseau L, Stankowski JN, Stetler C, Daughrity LM, Perkerson EA, Desaro P, Johnston A, Overstreet K, Edbauer D, Rademakers R, Boylan KB, Dickson DW, Fryer JD, Petrucelli L. Neurodegeneration. C9ORF72 repeat expansions in mice cause TDP-43 pathology, neuronal loss, and behavioral deficits. Science. 2015 Jun 5;348(6239):1151-4.

Busch JI, Unger TL, Jain N, Tyler Skrinak R, Charan RA, Chen-Plotkin AS. Increased expression of the frontotemporal dementia risk factor TMEM106B causes C9orf72-dependent alterations in lysosomes. Hum Mol Genet. 2016 Apr 28. pii: ddw127.

Cirulli ET, Lasseigne BN, Petrovski S, Sapp PC, Dion PA, Leblond CS, Couthouis J, Lu YF, Wang Q, Krueger BJ, Ren Z, Keebler J, Han Y, Levy SE, Boone BE, Wimbish JR, Waite LL, Jones AL, Carulli JP, Day-Williams AG, Staropoli JF, Xin WW, Chesi A, Raphael AR, McKenna-Yasek D, Cady J, Vianney de Jong JM, Kenna KP, Smith BN, Topp S, Miller J, Gkazi A; FALS Sequencing Consortium, Al-Chalabi A, van den Berg LH, Veldink J, Silani V, Ticozzi N, Shaw CE, Baloh RH, Appel S, Simpson E, Lagier-Tourenne C, Pulst SM, Gibson S, Trojanowski JQ, Elman L, McCluskey L, Grossman M, Shneider NA, Chung WK, Ravits JM, Glass JD, Sims KB, Van Deerlin VM, Maniatis T, Hayes SD, Ordureau A, Swarup S, Landers J, Baas F, Allen AS, Bedlack RS, Harper JW, Gitler AD, Rouleau GA, Brown R, Harms MB, Cooper GM, Harris T, Myers RM, Goldstein DB. Exome sequencing in amyotrophic lateral sclerosis identifies risk genes and pathways. Science. 2015 Mar 27;347(6229):1436-41.

Ciura S, Lattante S, Le Ber I, Latouche M, Tostivint H, Brice A, Kabashi E. Loss of function of C9orf72 causes motor deficits in a zebrafish model of Amyotrophic Lateral Sclerosis. Ann Neurol. 2013 May 30.

Dafinca R, Scaber J, Ababneh N, Lalic T, Weir G, Christian H, Vowles J, Douglas AG, Fletcher-Jones A, Browne C, Nakanishi M, Turner MR, Wade-Martins R, Cowley SA, Talbot K. C9orf72 Hexanucleotide Expansions are Associated with Altered ER Calcium Homeostasis and Stress Granule Formation in iPSC-Derived Neurons from Patients with

Amyotrophic Lateral Sclerosis and Frontotemporal Dementia. Stem Cells. 2016 Apr 20 doi: 10.1002/stem.2388. [Epub ahead of print]

DeJesus-Hernandez M, Mackenzie IR, Boeve BF, Boxer AL, Baker M, Rutherford NJ, Nicholson AM, Finch NA, Flynn H, Adamson J, Kouri N, Wojtas A, Sengdy P, Hsiung GY, Miller BL, Dickson DW, Boylan KB, Graff-Radford NR, Rademakers R. Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. Neuron. 2011 Oct 20;72(2):245-56.

Donnelly CJ, Zhang PW, Pham JT, Haeusler AR, Mistry NA, Vidensky S, Daley EL, Poth EM, Hoover B, Fines DM, Maragakis N, Tienari PJ, Petrucelli L, Traynor BJ, Wang J, Rigo F, Bennett CF, Blackshaw S, Sattler R, Rothstein JD. RNA toxicity from the ALS/FTD C9ORF72 expansion is mitigated by antisense intervention. Neuron. 2013 Oct 16;80(2):415-28.

Elden AC, Kim HJ, Hart MP, Chen-Plotkin AS, Johnson BS, Fang X, Armakola M, Geser F, Greene R, Lu MM, Padmanabhan A, Clay-Falcone D, McCluskey L, Elman L, Juhr D, Gruber PJ, Rüb U, Auburger G, Trojanowski JQ, Lee VM, Van Deerlin VM, Bonini NM, Gitler AD. Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS. Nature. 2010 Aug 26;466(7310):1069-75.

Farg MA, Sundaramoorthy V, Sultana JM, Yang S, Atkinson RA, Levina V, Halloran MA, Gleeson PA, Blair IP, Soo KY, King AE, Atkin JD. C9ORF72, implicated in amytrophic lateral sclerosis and frontotemporal dementia, regulates endosomal trafficking. Hum Mol Genet. 2014 Jul 1;23(13):3579-95.

Ferrari R, Mok K, Moreno JH, Cosentino S, Goldman J, Pietrini P, Mayeux R, Tierney MC, Kapogiannis D, Jicha GA, Murrell JR, Ghetti B, Wassermann EM, Grafman J, Hardy J, Huey ED, Momeni P. Screening for C9ORF72 repeat expansion in FTLD. Neurobiol Aging. 2012 Aug;33(8):1850.e1-11.

Freibaum BD, Lu Y, Lopez-Gonzalez R, Kim NC, Almeida S, Lee KH, Badders N, Valentine M, Miller BL, Wong PC, Petrucelli L, Kim HJ, Gao FB, Taylor JP. GGGGCC repeat expansion in C9orf72 compromises nucleocytoplasmic transport. Nature. 2015 Sep 3;525(7567):129-33.

Freischmidt A, Wieland T, Richter B, Ruf W, Schaeffer V, Müller K, Marroquin N, Nordin F, Hübers A, Weydt P, Pinto S, Press R, Millecamps S, Molko N, Bernard E, Desnuelle C, Soriani MH, Dorst J, Graf E, Nordström U, Feiler MS, Putz S, Boeckers TM, Meyer T, Winkler AS, Winkelman J, de Carvalho M, Thal DR, Otto M, Brännström T, Volk AE, Kursula P, Danzer KM, Lichtner P, Dikic I, Meitinger T, Ludolph AC, Strom TM, Andersen PM, Weishaupt JH. Haploinsufficiency of TBK1 causes familial ALS and fronto-temporal dementia. Nat Neurosci. 2015 May;18(5):631-6.

Gendron TF, Bieniek KF, Zhang YJ, Jansen-West K, Ash PE, Caulfield T, Daughrity L, Dunmore JH, Castanedes-Casey M, Chew J, Cosio DM, van Blitterswijk M, Lee WC, Rademakers R, Boylan KB, Dickson DW, Petrucelli L. Antisense transcripts of the expanded C9ORF72 hexanucleotide repeat form nuclear RNA foci and undergo repeat-associated non-ATG translation in c9FTD/ALS. Acta Neuropathol. 2013 Dec;126(6):829-44.

Haeusler AR, Donnelly CJ, Rothstein JD. The expanding biology of the C9orf72 nucleotide repeat expansion in neurodegenerative disease. Nat Rev Neurosci. 2016 Jun;17(6):383-95.

Haeusler AR, Donnelly CJ, Periz G, Simko EA, Shaw PG, Kim MS, Maragakis NJ, Troncoso JC, Pandey A, Sattler R, Rothstein JD, Wang J. C9orf72 nucleotide repeat structures initiate molecular cascades of disease. Nature. 2014 Mar 13;507(7491):195-200.

Hattula K, Peränen J. FIP-2, a coiled-coil protein, links Huntingtin to Rab8 and modulates cellular morphogenesis. Curr Biol. 2000 Dec 14-28;10(24):1603-6.

Heo JM, Ordureau A, Paulo JA, Rinehart J, Harper JW. The PINK1-PARKIN Mitochondrial Ubiquitylation Pathway Drives a Program of OPTN/NDP52 Recruitment and TBK1 Activation to Promote Mitophagy. Mol Cell. 2015 Oct 1;60(1):7-20.

Jiang J, Zhu Q, Gendron TF, Saberi S, McAlonis-Downes M, Seelman A, Stauffer JE, Jafar-Nejad P, Drenner K, Schulte D, Chun S, Sun S, Ling SC, Myers B, Engelhardt J, Katz M, Baughn M, Platoshyn O, Marsala M, Watt A, Heyser CJ, Ard MC, De Muynck L, Daughrity LM, Swing DA, Tessarollo L, Jung CJ, Delpoux A, Utzschneider DT, Hedrick SM, de Jong PJ, Edbauer D, Van Damme P, Petrucelli L, Shaw CE, Bennett CF, Da Cruz S, Ravits J, Rigo F, Cleveland DW, Lagier-Tourenne C. Gain of Toxicity from ALS/FTD-Linked Repeat Expansions in C9ORF72 Is Alleviated by Antisense Oligonucleotides Targeting GGGGCC-Containing RNAs. Neuron. 2016 May 4;90(3):535-50.

Jovičić A, Mertens J, Boeynaems S, Bogaert E, Chai N, Yamada SB, Paul JW 3rd, Sun S, Herdy JR, Bieri G, Kramer NJ, Gage FH, Van Den Bosch L, Robberecht W, Gitler AD. Modifiers of C9orf72 dipeptide repeat toxicity connect nucleocytoplasmic transport defects to FTD/ALS. Nat Neurosci. 2015 Sep;18(9):1226-9.

Koppers M, Blokhuis AM, Westeneng HJ, Terpstra ML, Zundel CA, Vieira de Sá R, Schellevis RD, Waite AJ, Blake DJ, Veldink JH, van den Berg LH, Jeroen Pasterkamp R. C9orf72 ablation in mice does not cause motor neuron degeneration or motor deficits. Ann Neurol. 2015 Jun 5. doi: 10.1002/ana.24453.

Lagier-Tourenne C, Baughn M, Rigo F, Sun S, Liu P, Li HR, Jiang J, Watt AT, Chun S, Katz M, Qiu J, Sun Y, Ling SC, Zhu Q, Polymenidou M, Drenner K, Artates JW, McAlonis-Downes M, Markmiller S, Hutt KR, Pizzo DP, Cady J, Harms MB, Baloh RH, Vandenberg SR, Yeo GW, Fu XD, Bennett CF, Cleveland DW, Ravits J. Targeted degradation of sense and antisense C9orf72 RNA foci as therapy for ALS and frontotemporal degeneration. Proc Natl Acad Sci U S A. 2013 Nov 19;110(47):E4530-9.

Lazarou M, Sliter DA, Kane LA, Sarraf SA, Wang C, Burman JL, Sideris DP, Fogel AI, Youle RJ. The ubiquitin kinase PINK1 recruits autophagy receptors to induce mitophagy. Nature. 2015 Aug 20;524(7565):309-14.

Levine TP, Daniels RD, Gatta AT, Wong LH, Hayes MJ. The product of C9orf72, a gene strongly implicated in neurodegeneration, is structurally related to DENN Rab-GEFs. Bioinformatics. 2013 Feb 15;29(4):499-503.

Liu Y, Pattamatta A, Zu T, Reid T, Bardhi O, Borchelt DR, Yachnis AT, Ranum LP. C9orf72 BAC Mouse Model with Motor Deficits and Neurodegenerative Features of ALS/FTD. Neuron. 2016 May 4;90(3):521-34.

Maruyama H, Morino H, Ito H, Izumi Y, Kato H, Watanabe Y, Kinoshita Y, Kamada M, Nodera H, Suzuki H, Komure O, A, Hirai T, Kato T, Ogasawara K, Hirano A, Takumi T, Kusaka H, Hagiwara K, Kaji R, Kawakami H. Mutations of optineurin in amyotrophic lateral sclerosis. Nature. 2010 May 13;465(7295):223-6.

Matsumoto G, Shimogori T, Hattori N, Nukina N. TBK1 controls autophagosomal engulfment of polyubiquitinated mitochondria through P62/SQSTM1 phosphorylation. Hum Mol Genet. 2015 May 13.

Mori K, Weng SM, Arzberger T, May S, Rentzsch K, Kremmer E, Schmid B, Kretzschmar HA, Cruts M, Van Broeckhoven C, Haass C, Edbauer D. The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTLD/ALS. Science. 2013 Mar 15;339(6125):1335-8.

Neumann M, Sampathu DM, Kwong LK, Truax AC, Micsenyi MC, Chou TT, Bruce J, Schuck T, Grossman M, Clark CM, McCluskey LF, Miller BL, Masliah E, Mackenzie IR, Feldman H, Feiden W, Kretzschmar HA, Trojanowski JQ, Lee VM. Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Science. 2006 Oct 6;314(5796):130-3.

O'Rourke JG, Bogdanik L, Muhammad AK, Gendron TF, Kim KJ, Austin A, Cady J, Liu EY, Zarrow J, Grant S, Ho R, Bell S, Carmona S, Simpkinson M, Lall D, Wu K, Daughrity L, Dickson DW, Harms MB, Petrucelli L, Lee EB, Lutz CM, Baloh RH. C9orf72 BAC Transgenic Mice Display Typical Pathologic Features of ALS/FTD. Neuron. 2015 Dec 2;88(5):892-901.

O'Rourke JG, Bogdanik L, Yáñez A, Lall D, Wolf AJ, Muhammad AK, Ho R, Carmona S, Vit JP, Zarrow J, Kim KJ, Bell S, Harms MB, Miller TM, Dangler CA, Underhill DM, Goodridge HS, Lutz CM, Baloh RH. C9orf72 is required for proper macrophage and microglial function in mice. Science. 2016 Mar 18;351(6279):1324-9.

Peters OM, Cabrera GT, Tran H, Gendron TF, McKeon JE, Metterville J, Weiss A, Wightman N, Salameh J, Kim J, Sun H, Boylan KB, Dickson D, Kennedy Z, Lin Z, Zhang YJ, Daughrity L, Jung C, Gao FB, Sapp PC, Horvitz HR, Bosco DA, Brown SP, de Jong P, Petrucelli L, Pilli M, Arko-Mensah J, Ponpuak M, Roberts E, Master S, Mandell MA, Dupont N, Ornatowski W, Jiang S, Bradfute SB, Bruun JA, Hansen TE, Johansen T, Deretic V. TBK-1 promotes autophagy-mediated antimicrobial defense by controlling autophagosome maturation. Immunity. 2012 Aug 24;37(2):223-34.

Renton AE, Majounie E, Waite A, Simón-Sánchez J, Rollinson S, Gibbs JR, Schymick JC, Laaksovirta H, van Swieten JC, Myllykangas L, Kalimo H, Paetau A, Abramzon Y, Remes AM, Kaganovich A, Scholz SW, Duckworth J, Ding J, Harmer DW, Hernandez DG, Johnson JO, Mok K, Ryten M, TRabzuni D, Guerreiro RJ, Orrell RW, Neal J, Murray A, Pearson J, Jansen IE, Sondervan D, Seelaar H, Rothstein JD, Sendtner M, Drepper C, Eichler EE, Alkan C, Abdullaev Z, Pack SD, Dutra A, Pak E, Hardy J, Singleton A, Williams NM, Heutink P, Pickering-Brown S, Morris HR, Tienari

PJ, Traynor BJ. A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. Neuron. 2011 Oct 20;72(2):257-68.

Sato T, Iwano T, Kunii M, Matsuda S, Mizuguchi R, Jung Y, Hagiwara H, Yoshihara Y, Yuzaki M, Harada R, Harada A. Rab8a and Rab8b are essential for several apical transport pathways but insufficient for ciliogenesis. J Cell Sci. 2014 Jan 15;127(Pt 2):422-31.

Sellier C, Campanari ML, Julie Corbier C, Gaucherot A, Kolb-Cheynel I, Oulad-Abdelghani M, Ruffenach F, Page A, Ciura S, Kabashi E, Charlet-Berguerand N. Loss of C9ORF72 impairs autophagy and synergizes with polyQ Ataxin-2 to induce motor neuron dysfunction and cell death. EMBO J. 2016 Apr 21. pii: e201593350. [Epub ahead of print]

Seto S, Sugaya K, Tsujimura K, Nagata T, Horii T, Koide Y. Rab39a interacts with phosphatidylinositol 3-kinase and negatively regulates autophagy induced by lipopolysaccharide stimulation in macrophages. PLoS One. 2013 Dec 13;8(12):e83324.

Sullivan PM, Zhou X, Robins AM, Paushter DH, Kim D, Smolka MB, Hu F. The ALS/FTLD associated protein C9orf72 associates with SMCR8 and WDR41 to regulate the autophagy-lysosome pathway. Acta Neuropathol Commun. 2016 May 18;4(1):51.

Tao Z, Wang H, Xia Q, Li K, Li K, Jiang X, Xu G, Wang G, Ying Z. Nucleolar stress and impaired stress granule formation contribute to C9orf72 RAN translation-induced cytotoxicity. Hum Mol Genet. 2015 May 1;24(9):2426-41.

Van Blitterswijk M, van Es MA, Hennekam EA, Dooijes D, van Rheenen W, Medic J, Bourque PR, Schelhaas HJ, van der Kooi AJ, de Visser M, de Bakker PI, Veldink JH, van den Berg LH. Evidence for an oligogenic basis of amyotrophic lateral sclerosis. Hum Mol Genet. 2012 Sep 1;21(17):3776-84.

van Blitterswijk M, Gendron TF, Baker MC, DeJesus-Hernandez M, Finch NA, Brown PH, Daughrity LM, Murray ME, Heckman MG, Jiang J, Lagier-Tourenne C, Edbauer D, Cleveland DW, Josephs KA, Parisi JE, Knopman DS, Petersen RC, Petrucelli L, Boeve BF, Graff-Radford NR, Boylan KB, Dickson DW, Rademakers R. Novel clinical associations with specific C9ORF72 transcripts in patients with repeat expansions in C9ORF72. Acta Neuropathol. 2015 Dec;130(6):863-76.

Waite AJ, Bäumer D, East S, Neal J, Morris HR, Ansorge O, Blake DJ. Reduced C9orf72 protein levels in frontal cortex of amyotrophic lateral sclerosis and frontotemporal degeneration brain with the C9ORF72 hexanucleotide repeat expansion. Neurobiol Aging. 2014 Jul;35(7):1779.e5-1779.e13.

Webster CP, Smith EF, Bauer CS, Moller A, Hautbergue GM, Ferraiuolo L, Myszczynska MA, Higginbottom A, Walsh MJ, Whitworth AJ, Kaspar BK, Meyer K, Shaw PJ, Grierson AJ, De Vos KJ. The C9orf72 protein interacts with Rab1a and the ULK1 complex to regulate initiation of autophagy. EMBO J. 2016 Jun 22. pii: e201694401. [Epub ahead of print]

Xiao S, MacNair L, McLean J, McGoldrick P, McKeever P, Soleimani S, Keith J, Zinman L, Rogaeva E, Robertson J. C9orf72 isoforms in Amyotrophic Lateral Sclerosis and Frontotemporal Lobar Degeneration. Brain Res. 2016 Apr 29.

Zhang D, Iyer LM, He F, Aravind L. Discovery of Novel DENN Proteins: Implications for the Evolution of Eukaryotic Intracellular Membrane Structures and Human Disease. Front Genet. 2012 Dec 13;3:283.

Zhang K, Donnelly CJ, Haeusler AR, Grima JC, Machamer JB, Steinwald P, Daley EL, Miller SJ, Cunningham KM, Vidensky S, Gupta S, Thomas MA, Hong I, Chiu SL, Huganir RL, Ostrow LW, Matunis MJ, Wang J, Sattler R, Lloyd TE, Rothstein JD. The C9orf72 repeat expansion disrupts nucleocytoplasmic transport. Nature. 2015 Sep 3;525(7567):56-61.

Zu T, Gibbens B, Doty NS, Gomes-Pereira M, Huguet A, Stone MD, Margolis J, Peterson M, Markowski TW, Ingram MA, Nan Z, Forster C, Low WC, Schoser B, Somia NV, Clark HB, Schmechel S, Bitterman PB, Gourdon G, Swanson MS, Moseley M, Ranum LP. Non-ATG-initiated translation directed by microsatellite expansions. Proc Natl Acad Sci U S A. 2011 Jan 4;108(1):260-5.

Zu T, Liu Y, Bañez-Coronel M, Reid T, Pletnikova O, Lewis J, Miller TM, Harms MB, Falchook AE, Subramony SH, Ostrow LW, Rothstein JD, Troncoso JC, Ranum LP. RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. Proc Natl Acad Sci U S A. 2013 Dec 17;110(51):E4968-77.

Figure 1

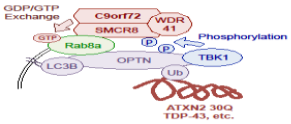


Figure 1. Tentative model of C9ORF72 role in autophagy. C9ORF72 forms a complex with the SMCR8 and WDR41 proteins. This complex acts as a GDP/GTP exchange factor for the small GTPases Rab8a that interacts with the autophagy receptor Optineurin (OPTN), or for Rab39b that interacts with P62, an autophagy receptors alike to OPTN. SMCR8 is phosphorylated and potentially activated by the TBK1 kinase, which also interacts with OPTN. Optineurin bridges ubiquitinated proteins to the autophagosome membrane through the LC3B proteins. Recruitment of Rab8a and TBK1 by OPTN allows initiation of autophagy at the precise site of protein aggregate, dysfunctional organelles or intracellular pathogen. However, it remains to determine the function of Rab8a or RAb39b in autophagy, notably whether Rab8a or Rab39b may promote autophagosome membrane elongation or fusion.

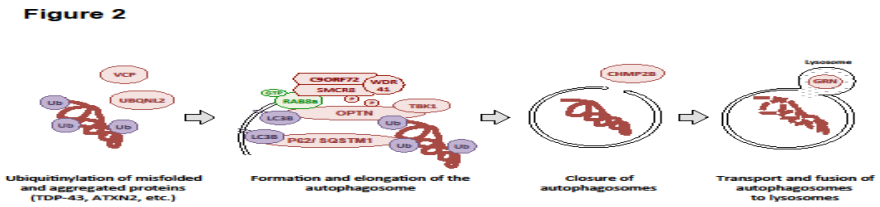


Figure 2. Various genes muted in ALS-FTD regulate endo-lysosomal and protein degradation pathways. The kinase TBK1 activates autophagy through phosphorylation of the autophagy receptors P62 and OPTN. P62, encoded by the SQSTM1 gene, and OPTN bridge ubiquitinated proteins to be degraded with lipidated LC3B proteins that are bound to the autophagy membrane. The Valosin-Containing protein (VCP) is an ATPase that unfolds substrate ubiquitinated protein complex allowing the released proteins to be degraded by the proteasome or by autophagy. The ubiquitin-like protein Ubiquilin-2 (UBQLN2) functionally links the ubiquitination machinery to the proteasome and autophagy machineries. Charged multivesicular body protein 2b (CHMP2B) belongs to the Endosomal Sorting Complex III (ESCRT-III) required for closure of membrane vesicles and notably of multivesicular bodies and autophagosomes. Finally, progranulin (GRN) is localized to the lysosome, which is the journey's end of endosomes and autophagosomes. Mutation in TBK1, SQSTM1, OPTN, UBQLN2, VCP cause ALS-FTD, while haploinsufficiency mutations of GRN causes FTD.

AUTOPHAGY
2016, VOL. 12, NO. 8, 1406–1408
http://dx.doi.org/10.1080/15548627.2016.1189070

Taylor & Francis
Taylor & Francis Group

AUTOPHAGIC PUNCTUM

OPEN ACCESS

The most prevalent genetic cause of ALS-FTD, C9orf72 synergizes the toxicity of ATXN2 intermediate polyglutamine repeats through the autophagy pathway

Sorana Ciura[a], Chantal Sellier[b], Maria-Letizia Campanari[a], Nicolas Charlet-Berguerand[b], and Edor Kabashi[a]

[a]Sorbonne Université, Université Pierre et Marie Curie (UPMC), Université de Paris 06, Unité Mixte 75, Institut National de la Santé et de la Recherche Médicale (INSERM) Unité 1127, Centre National de la Recherche Scientifique (CNRS) Unité Mixte de Recherche 7225 Institut du Cerveau et de la Moelle Épinière (ICM), Paris, France; [b]Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), INSERM U964, CNRS UMR7104, Strasbourg University, Illkirch, France

ABSTRACT

The most common genetic cause for amyotrophic lateral sclerosis and frontotemporal dementia (ALS-FTD) is repeat expansion of a hexanucleotide sequence (GGGGCC) within the *C9orf72* genomic sequence. To elucidate the functional role of C9orf72 in disease pathogenesis, we identified certain molecular interactors of this factor. We determined that C9orf72 exists in a complex with SMCR8 and WDR41 and that this complex acts as a GDP/GTP exchange factor for RAB8 and RAB39, 2 RAB GTPases involved in macroautophagy/autophagy. Consequently, C9orf72 depletion in neuronal cultures leads to accumulation of unresolved aggregates of SQSTM1/p62 and phosphorylated TARDBP/TDP-43. However, C9orf72 reduction does not lead to major neuronal toxicity, suggesting that a second stress may be required to induce neuronal cell death. An intermediate size of polyglutamine repeats within ATXN2 is an important genetic modifier of ALS-FTD. We found that coexpression of intermediate polyglutamine repeats (30Q) of ATXN2 combined with C9orf72 depletion increases the aggregation of ATXN2 and neuronal toxicity. These results were confirmed in zebrafish embryos where partial C9orf72 knockdown along with intermediate (but not normal) repeat expansions in ATXN2 causes locomotion deficits and abnormal axonal projections from spinal motor neurons. These results demonstrate that C9orf72 plays an important role in the autophagy pathway while genetically interacting with another major genetic risk factor, ATXN2, to contribute to ALS-FTD pathogenesis.

ARTICLE HISTORY
Received 2 May 2016
Revised 4 May 2016
Accepted 6 May 2016

KEYWORDS
amytrophic lateral sclerosis (ALS); ATXN2 (ataxin 2); C9orf72; frontotemporal dementia (FTD); neurodegeneration; SQSTM1/p62; TBK1; TARDBP/TDP-43; zebrafish

Amyotrophic lateral sclerosis (ALS), a neurodegenerative disease affecting motor neurons, is the third most common neurodegenerative disease affecting one in 50,000 people. Frontotemporal dementia (FTD), which affects neurons from the frontal and temporal cortex, is the second most common presenile dementia after Alzheimer disease. An expansion of hundreds to thousands of GGGGCC repeats within the first intron of the *C9orf72* gene represents the most common inherited cause for ALS and FTD accounting for 20–60% of familial patients and 1–7% of sporadic patients for these disorders in Northern Europe and North America. Major effort has been dedicated over the past 5 y to unravel the mechanism of toxicity through which C9orf72 repeat expansion leads to neurodegeneration. Notably, noncanonical translation, termed repeat-associated non-ATG (RAN) translation, of these expanded GGGGCC repeats generates repeated dipeptide chains that are toxic in neuronal cultures as well as in yeast, *Drosophila*, and mice.

In parallel, pathologically expanded GGGGCC repeats were shown to affect C9orf72 expression with reduced transcript and protein levels consistently measured in pathological tissue obtained from ALS and FTD patients. Therefore, C9orf72 haploinsufficiency may also participate in neuronal degeneration in ALS-FTD. In line with this evidence, reduced levels of C9orf72 in *C. elegans* and zebrafish, but not in mice, trigger specific phenotypic features associated with motor neuron degeneration. Importantly, very little is yet known about the role of C9orf72, and a better comprehension of its function will be essential to understand its implication in ALS-FTD. To ascertain the molecular and cellular function for C9orf72, we applied an initial proteomic approach to identify its partners. Tandem tag immunoprecipitation revealed that C9orf72 exists in a complex with 2 proteins, SMCR8 and WDR41. These results were confirmed by co-immunoprecipitation experiments and by reconstituting this complex with recombinant proteins from baculovirus-infected insect cells. Both SMCR8 and C9orf72 contain DENN domains, which are typical of GDP/GTP exchange factors (GEFs) for small RAB GTPases. Indeed, in vitro the complex formed by C9orf72, SMCR8 and WDR41 promotes

CONTACT Nicolas Charlet-Berguerand ncharlet@igbmc.fr IGBMC, 1 Rue Laurent Fries, 67404 Illkirch, France; Edor Kabashi edor.kabashi@icm-institute.org ICM, 47 Boulevard de l'Hôpital, Paris France 75013

Punctum to: Sellier C, Campanari ML, Julie Corbier C, Gaucherot A, Kolb-Cheynel I, Oulad-Abdelghani M, Ruffenach F, Page A, Ciura S, Kabashi E, Charlet-Berguerand N. Loss of C9ORF72 impairs autophagy and synergizes with polyQ Ataxin-2 to induce motor neuron dysfunction and cell death. EMBO J. 2016; pii: e201593350.

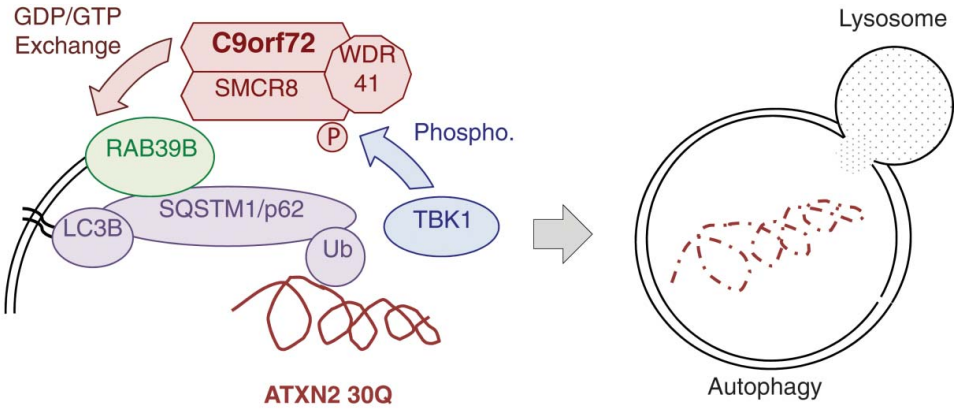


Figure 1. Tentative model of C9orf72 function. C9orf72 forms a complex with the SMCR8 and WDR41 proteins and acts as a GDP/GTP exchange factor for the small RAB GTPase RAB39B. SMCR8 is phosphorylated and potentially activated by the TBK1 kinase. The partially reduced expression of C9orf72 partly impairs autophagy but does not cause massive neuronal cell death. In contrast, reduced expression of C9orf72 synergizes the aggregation and toxicity of ATXN2 with intermediate length polyglutamine repeats. These results suggest a double-hit hypothesis for ALS-FTD.

GDP/GTP exchange for the small GTPases RAB8A and RAB39B. Since RAB8 and RAB39 are involved in macroautophagy and since SMCR8 was found in proteomic analysis of the autophagy network, we investigated whether C9orf72 was regulating macroautophagy. Knockdown of C9orf72 by shRNA and/or siRNA in neuronal primary cultures derived from mouse frontal cortex demonstrate that a partial depletion of C9orf72 has a deleterious effect on autophagy, with notable accumulation of unresolved aggregates of the autophagy receptor SQSTM1. In contrast, the deleterious effect of the loss of C9orf72 on the lipidation of LC3B was rather mild. Also, we found that the C9orf72 complex associates with the autophagy receptors OPTN and SQSTM1, probably through indirect interaction via RAB8 or RAB39 or other proteins, yet to be identified. Note that mutations in OPTN and SQSTM1 cause ALS-FTD and that SQSTM1-positive aggregates are observed in ALS-FTD patients with expansion of GGGGCC repeats in C9orf72. The role of the C9orf72 complex in autophagy was further strengthened as 2 kinases regulating autophagy, the ULK1 and TBK1 kinases, were shown to phosphorylate SMCR8. The TBK1 phosphorylation of SMCR8 is important to neuronal cells as expression of a SMCR8 form mimicking TBK1-constitutive phosphorylation restores the autophagy imbalance due to depletion of SMCR8 or TBK1. Again, the link with ALS is patent as loss-of-function mutations in TBK1 were recently shown to cause ALS-FTD.

Importantly, the autophagy deficiency observed in neuronal cultures upon knockdown of C9orf72 or TBK1 can be rescued upon co-expression of a constitutively active GTP-locked form of RAB39B that does not require GEF activity. Of interest, loss-of-function mutations of RAB39B cause mental retardation associated with early onset Parkinson disease. Therefore, our results indicate that interaction of C9orf72 with TBK1-phosphorylated SMCR8 and their association and consequent activation of RAB39B has an important role for autophagy in neurons as shown in Figure 1.

Furthermore, our data indicate that decreased expression of C9orf72 leads to some accumulation of TARDBP aggregates, which is a pathological hallmark of ALS-FTD. Therefore, reduced expression of C9orf72 in neuronal cultures recapitulates some of the pathological features of C9orf72-associated ALS-FTD. However, autophagy deregulation by C9orf72 loss of function was not sufficient to trigger neuronal cell death in mammalian neuronal cultures. Therefore, we hypothesized that an extra cellular stressor was required to trigger neurodegeneration upon C9orf72 loss of function.

ATXN2 (ataxin 2) with intermediate polyglutamine (polyQ) repeats represents the most important risk factor for ALS-FTD, and we have recently demonstrated that C9orf72 repeat expansions coincide with intermediate size of polyQ in ALS-FTD patients. Of interest to disease pathogenesis, partial C9orf72 depletion coupled with intermediate polyQ repeats (30x) leads to increased ATXN2 aggregation and neuronal cell death. This epistatic interaction appears specific for ATXN2 intermediate polyQ repeats, because C9orf72 loss of function does not accentuate toxicity of other proteins harboring polyQ repeats (ATXN3, HTT) or other ALS-related mutants (FUS and SOD1), at least in neuronal culture and in the time frame of our study. Importantly, the genetic interaction described in neuronal cultures was confirmed in vivo in zebrafish embryos where partial C9orf72 knockdown coupled with intermediate (30x), but not normal-size (22x), polyQ repeats of ATXN2 is associated with swimming deficits and aberrant axonal projections from spinal motor neurons (Fig. 1).

The exact regulatory role of C9orf72 complex in the autophagy pathway and the molecular mechanism through which C9orf72 partial loss of function accentuates the toxicity of ATXN2 intermediate repeats remain to be established. Similarly, the consequences of SMCR8 phosphorylation by ULK1 and TBK1 or other kinases remain to be ascertained. Also, the synergic toxicity observed could be caused by independent molecular cascades that converge on protein degradation through the autophagosome. In the model that we propose, increased polyQ size renders ATXN2 more stable and potentially less available to degradation. This is aggravated

by C9orf72 depletion through a deregulation of the autophagosome capacity. Overall, our study provides a better understanding of the function of C9orf72, therefore opening new perspectives on its potential role in ALS-FTD pathogenesis.

Abbreviations

ATXN2	ataxin 2
C9orf72	chromosome 9 open reading frame 72
polyQ	polyglutamine
SQSTM1/p62	sequestosome 1
SMCR8	Smith-Magenis syndrome chromosome region, candidate 8 homolog
TARDBP/TDP-43	TAR DNA binding protein
TBK1	tank binding kinase 1
WDR41	WD repeat domain 41

Disclosure of potential conflicts of interest

No potential conflicts of interest were disclosed.

Funding

This work was supported by Fondation de France Thierry Latran #57486 "Model-ALS," AFM grant #18605 "Role of C9ORF72 in ALS-FTD," ERC-2012-StG #310659 "RNA DISEASES," ANR-10-LABX-0030-INRT and ANR-10-IDEX-0002–02 (NCB); Atip/Avenir from Inserm, Career Integration Grant (Marie Curie Actions), Robert Packard Foundation, E-rare ERA-NET program, AFM, ARSLA, France-Alzheimer association and the program "Investissements d'avenir" ANR-10-IAIHU-06 (EK).

Article

SOURCE DATA | TRANSPARENT PROCESS

THE EMBO JOURNAL

Loss of C9ORF72 impairs autophagy and synergizes with polyQ Ataxin-2 to induce motor neuron dysfunction and cell death

Chantal Sellier[1,*], Maria-Letizia Campanari[2], Camille Julie Corbier[1], Angeline Gaucherot[1], Isabelle Kolb-Cheynel[1], Mustapha Oulad-Abdelghani[1], Frank Ruffenach[1], Adeline Page[1], Sorana Ciura[2], Edor Kabashi[2] & Nicolas Charlet-Berguerand[1,**]

Abstract

An intronic expansion of GGGGCC repeats within the C9ORF72 gene is the most common genetic cause of amyotrophic lateral sclerosis and frontotemporal dementia (ALS-FTD). Ataxin-2 with intermediate length of polyglutamine expansions (Ataxin-2 Q30x) is a genetic modifier of the disease. Here, we found that C9ORF72 forms a complex with the WDR41 and SMCR8 proteins to act as a GDP/GTP exchange factor for RAB8a and RAB39b and to thereby control autophagic flux. Depletion of C9orf72 in neurons partly impairs autophagy and leads to accumulation of aggregates of TDP-43 and P62 proteins, which are histopathological hallmarks of ALS-FTD. SMCR8 is phosphorylated by TBK1 and depletion of TBK1 can be rescued by phosphomimetic mutants of SMCR8 or by constitutively active RAB39b, suggesting that TBK1, SMCR8, C9ORF72, and RAB39b belong to a common pathway regulating autophagy. While depletion of C9ORF72 only has a partial deleterious effect on neuron survival, it synergizes with Ataxin-2 Q30x toxicity to induce motor neuron dysfunction and neuronal cell death. These results indicate that partial loss of function of C9ORF72 is not deleterious by itself but synergizes with Ataxin-2 toxicity, suggesting a double-hit pathological mechanism in ALS-FTD.

Keywords C9ORF72; autophagy; neurodegeneration; ALS-FTD
Subject Categories Neuroscience
DOI 10.15252/embj.201593350 | Received 23 October 2015 | Revised 14 March 2016 | Accepted 15 March 2016 | Published online 21 April 2016
The EMBO Journal (2016) 35: 1276–1297

See also: **S Almeida & F-B Gao** (June 2016)

Introduction

Amyotrophic lateral sclerosis (ALS), a motor neuron degenerative disease, and frontotemporal dementia (FTD), a presenile dementia affecting frontal and temporal brain regions, share clinical, genetic, and pathological overlap and are now considered in some cases as manifestations of a similar disease continuum (Lomen-Hoerth *et al*, 2002; Ringholz *et al*, 2005; Neumann *et al*, 2006). A notion that is emphasized by the identification of expanded GGGGCC repeats within the first intron of the *C9ORF72* gene as the most common inherited cause for both ALS and FTD (DeJesus-Hernandez *et al*, 2011; Renton *et al*, 2011; Gijselinck *et al*, 2012; Majounie *et al*, 2012).

Three non-exclusive mechanisms by which expanded GGGGCC repeats cause neuron degeneration have been proposed. First, sense and antisense transcripts containing expanded GGGGCC repeats accumulate in nuclear RNA aggregates, which recruit specific RNA-binding proteins, thereby potentially inhibiting their functions (Almeida *et al*, 2013; Donnelly *et al*, 2013; Lagier-Tourenne *et al*, 2013; Mizielinska *et al*, 2013). Various proteins have been reported to bind to GGGGCC RNA repeats, but their roles in pathogenesis are, yet, to be determined (Lee *et al*, 2013; Mori *et al*, 2013a; Cooper-Knock *et al*, 2014; Haeusler *et al*, 2014). The second potential mechanism for neurotoxicity of CCGGGG expansions is a form of non-canonical protein translation termed repeat-associated non-ATG (RAN) translation (Zu *et al*, 2013). Indeed, expanded GGGGCC repeats are RAN-translated in all six sense and antisense frames, resulting in expression of five different dipeptide repeats containing proteins (DPRs or DRPs, also named C9RANT), which form inclusions throughout the brain of patients with C9-ALS/FTD (Ash *et al*, 2013; Gendron *et al*, 2013; Mori *et al*, 2013b; Zu *et al*, 2013), as well as in mice expressing expanded CCGGGG repeats (Chew *et al*, 2015; O'Rourke *et al*, 2015; Peters *et al*, 2015). These DPRs were recently shown to be toxic in neuronal cell cultures and in *Drosophila* models through alteration of the nucleocytoplasmic transport (Kwon *et al*, 2014; May *et al*, 2014; Mizielinska *et al*, 2014; Wen

1 Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), INSERM U964, CNRS UMR7104, Strasbourg University, Illkirch, France
2 Sorbonne Université, Université Pierre et Marie Curie (UPMC), Université de Paris 06, Unité Mixte 75, Institut National de la Santé et de la Recherche Médicale (INSERM) Unité 1127, Centre National de la Recherche Scientifique (CNRS) Unité Mixte de Recherche 7225, Institut du Cerveau et de la Moelle Épinière (ICM), 75013 Paris, France
*Corresponding author. Tel: +33 388 653 309; Fax: +33 388 653 201; E-mail: sellier@igbmc.fr
**Corresponding author. Tel: +33 388 653 309; Fax: +33 388 653 201; E-mail: ncharlet@igbmc.fr

et al, 2014; Zhang et al, 2014, 2015; Freibaum et al, 2015; Jovičić et al, 2015; Tao et al, 2015). Third, decreased expression of *C9ORF72* mRNA expression levels in C9-ALS/FTD patients (DeJesus-Hernandez et al, 2011; Gijselinck et al, 2012; Almeida et al, 2013; Waite et al, 2014; van Blitterswijk et al, 2015), as well as motor deficit caused by knockdown of C9orf72 in zebrafish (Ciura et al, 2013), suggests that haploinsufficiency of C9ORF72 may participate in neuronal degeneration. However, the absence of neuronal phenotypes in mouse depleted of *C9orf72* expression in brain or in neurons (Lagier-Tourenne et al, 2013; Koppers et al, 2015), as well as the absence of ALS/FTD patients with null alleles or missense mutations in *C9ORF72*, strongly argues against a sole loss of function of C9ORF72 as the cause of ALS-FTD.

Despite these advances, little is known about the normal molecular and cellular functions of C9ORF72. Bioinformatics analysis predicts that C9ORF72 contains DENN (Differentially Expressed in Normal and Neoplastic cells) domains characteristic of Rab GDP/GTP exchange factors (GEFs) (Levine et al, 2013; Zhang et al, 2012). Rab GTPases are monomeric guanine nucleotide-binding proteins that switch between two conformational states, an inactive form bound to GDP and an active form bound to GTP. GEF proteins catalyze the conversion of Rab proteins from GDP-bound to GTP-bound form, thereby activating Rab functions. Rab GTPases regulate many steps of membrane traffic, including vesicle formation, vesicle movement, and membrane fusion. Consistent with a function of C9ORF72 as a GEF protein, C9ORF72 was reported to interact with Rab1, Rab5, Rab7, and Rab11 and its depletions leads to endocytosis and autophagy dysfunctions (Farg et al, 2014).

Macroautophagy, named autophagy thereafter, is a catabolic process that engulfs cytoplasmic constituents within a double-membrane vesicle named autophagosome, which is directed to the lysosome for degradation and recycling. Hence, autophagy plays an essential role in homeostasis by providing energy and recycling cellular components, but also by facilitating lysosomal degradation of intracellular pathogens, defective organelles, and aggregates of misfolded proteins. In that aspect, autophagy is crucial for normal brain function (Hara et al, 2006; Komatsu et al, 2006), and autophagy dysfunction has been reported in various neurodegenerative diseases, including Alzheimer's, Parkinson's, and Huntington's diseases, as well as ALS and FTD (review in Wong & Cuervo, 2010; Nixon, 2013).

A better comprehension of C9ORF72 functions is essential to understand its possible implication in ALS-FTD. Here, we found that C9ORF72 forms a complex with WDR41 and SMCR8, a DENN protein previously identified within the interactome of autophagy (Behrends et al, 2010). We found that the complex formed by C9ORF72, SMCR8, and WDR41 interacts and acts as a GDP/GTP exchange factor for RAB8a and RAB39b, which are Rab GTPases involved in vesicle trafficking and autophagy (Pilli et al, 2012; Seto et al, 2013; Sato et al, 2014).

Next, we found that decreased expression of C9ORF72 partly inhibits autophagy and leads to accumulation of cytoplasmic aggregates of P62/SQSTM1 and of TDP-43, thus recapitulating two histopathological hallmarks of ALS-FTD patients (Neumann et al, 2006; Al-Sarraj et al, 2011). Also, SMCR8, but not C9ORF72 or WDR41, is phosphorylated by ULK1 or TBK1, which are kinases regulating autophagy (Chan et al, 2007; Hara et al, 2008; Thurston et al, 2009; Wild et al, 2011; Pilli et al, 2012; Matsumoto et al, 2015; Heo et al, 2015; Lazarou et al, 2015). Importantly, either a mutant of SMCR8 mimicking a constitutive TBK1-dependent phosphorylation or a constitutively active mutant of RAB39b that does not require GEF activity can correct alteration of autophagy caused by depletion of TBK1 or C9ORF72. These results suggest that TBK1, C9ORF72 complex, and RAB39b belong to a common pathway regulating autophagy in neurons.

Finally, we found that decreased expression of C9ORF72 potentiates the aggregation and toxicity of Ataxin-2 with intermediate length of polyglutamine expansions (Ataxin-2 Q30x) but not of Ataxin-2 with normal polyQ length (Ataxin-2 Q22x). This is relevant as intermediate size of polyglutamine expansions in Ataxin-2 is a genetic modifier of ALS-FTD (Elden et al, 2010; Daoud et al, 2011; Ross et al, 2011; Van Damme et al, 2011; Lattante et al, 2014).

In conclusion, these results provide a molecular and cellular function for C9ORF72 as a GEF protein regulating autophagy, but also support a double-hit mechanism in ALS-FTD, where the sole haploinsufficiency of *C9ORF72* might not be sufficient to cause neuronal cell death but may synergize the neurodegeneration caused by accumulation of toxic proteins, including Ataxin-2 with intermediate polyQ length.

Results

C9ORF72 forms a complex with the WDR41 and SMCR8 proteins

To better characterize the function of C9ORF72, we performed a proteomic analysis to identify potential interactants of C9ORF72. Of technical interest, transfection of Flag-tagged human C9ORF72 cDNA resulted in low expression of C9ORF72 protein. Bioinformatics analysis indicated that mRNA sequences of human or mouse *C9ORF72* present an excess of rare codons and negative RNA *cis* element that would impair expression of C9ORF72. Thus, we cloned an optimized sequence of human *C9ORF72* cDNA expressing the exact same amino acid sequence, but in which codon usage was optimized and negative RNA *cis* elements were removed. Neuronal N2A mouse cells were transfected with HA-Flag-tagged optimized C9ORF72 and tandem-tag purification followed by nano-LC-MS/MS analysis identified various proteins, including Rab8a, Rab39b, Smcr8, Wdr41, P62, Hsc70, Hsp90, and Bag3 (Fig 1A and Table EV1). Bioinformatics analysis of this putative C9ORF72 interactome using the DAVID database (NIAID, NIH) predicted a potential association with ALS disease (P-value of 4.3×10^{-2}), and analysis of KEGG, GO, and Reactome biological pathways using Gene Set Enrichment Analysis (GSEA, Broad institute) revealed significant enrichment for adaptive immune system (FDR q-value, 2.6×10^{-6}), activation of NF-κB (FDR q-value, 9.8×10^{-5}), formation of phagosome and autophagosome (FDR q-values of 8.8×10^{-5} and 1.4×10^{-4}, respectively), vesicle-mediated transport (FDR q-value, 6.7×10^{-3}), and proteasome (FDR q-value, 6.3×10^{-3}).

Western blotting on tandem-tag-purified proteins confirmed association of endogenous Rab8a, Rab39b, Smcr8, Wdr41, P62, Hsc70, and Bag3 with transfected HA-Flag-tagged C9ORF72 (Fig EV1A). RAB8 and RAB39 are Rab GTPases involved in vesicle trafficking and autophagy (Pilli et al, 2012; Seto et al, 2013; Sato et al, 2014). WDR41 is a 52-kDa protein of unknown function that contains a protein–protein interaction domain consisting of six WD40 repeats.

Figure 1. C9ORF72 in complex with SMCR8 and WDR41 is a GEF for RAB8 and RAB39.

A Silver staining of proteins extracted from N2A mouse neuronal cells expressing Flag-HA-tagged C9ORF72 and captured through consecutive anti-Flag and anti-HA affinity purification steps.

B Immunoblot analysis of HA-immunoprecipitated proteins and lysate of HEK293 cells co-expressing HA-tagged C9ORF72 and/or Flag-tagged SMCR8 and/or Flag-tagged WDR41.

C Coomassie blue staining of Nickel-NTA affinity purification of HIS-C9ORF72, SMCR8, and WDR41 co-expressed in baculovirus-infected insect cells.

D Immunoblot analysis of HA-immunoprecipitated proteins and lysate of HEK293 cells co-expressing HA-tagged C9ORF72 and HA-tagged SMCR8 with various Flag-tagged Rab GTPases.

E Immunoblot against endogenous C9orf72, Wdr41, Rab8a, Rab39b, Rab5, and Rab7 of control (IgG alone) or endogenous Smcr8 immunoprecipitated from adult mouse brain.

F α-^{32}P-radiolabelled GDP release from GST-tagged purified RAB8a as a function of increased concentration of either recombinant purified C9ORF72 alone or in complex with SMCR8 and WDR41.

G Identical GDP release assay as in (F) but using recombinant purified GST-tagged RAB39b instead of RAB8a.

H Schematic representation of the C9ORF72 complex acting as a Rab-guanine nucleotide exchange factor.

Data information: Error bars indicate SEM. Experiments were repeated 3 times (n = 3).
Source data are available online for this figure.

Smith–Magenis syndrome chromosome region candidate 8 (SMCR8) is a 105-kDa protein that, similar to C9ORF72, contains a divergent Differentially Expressed in Normal and Neoplasia (DENN) domain (Zhang *et al*, 2012; Levine *et al*, 2013). Interestingly, SMCR8 has been found to interact with the ULK1 kinase complex, which initiates autophagy (Behrends *et al*, 2010). P62, encoded by the *SQSTM1* gene, binds to poly-ubiquitinated proteins and targets these proteins for autophagy degradation. Finally, HSC70, encoded by the *HSPA8* gene, is a constitutively expressed heat-shock protein that among various functions interacts with BAG3 (BCL2-associated athanogene 3) and the ubiquitin ligase STUB1 (also known as CHIP) to initiate aggrephagy, also named chaperone-assisted selective autophagy (CASA), which is the selective autophagy of chaperone-bound misfolded or aggregated proteins (Gamerdinger *et al*, 2009; Arndt *et al*, 2010). In contrast, we found no interaction of C9ORF72 with other putative candidates identified by proteomic analysis, such as Senataxin, TDP-43, IQGAP, and the proteasome proteins PSMD4, PSMD8, PSMD10, while some other proteins were not tested (DCTN, NSF, SNX, STX17, VPS16, SEC22, etc.).

Next, we confirmed by cell transfection and co-immunoprecipitation experiments that HA-tagged C9ORF72 interacts with Flag-tagged SMCR8 and that the presence of both SMCR8 and C9ORF72 is required to recruit Flag-tagged WDR41 (Fig 1B). The association of C9ORF72 with SMCR8 and WDR41 is direct since HIS-tagged C9ORF72 was able to pull down both recombinant SMCR8 and WDR41 expressed in baculovirus-infected insect cells (Fig 1C). Of technical interest, we noted that the interaction of C9ORF72 with SMCR8 stabilizes and increases expression of both C9ORF72 and SMCR8. *C9ORF72* also encodes a putative shorter splicing form of ~30 kDa, but co-immunoprecipitation assays demonstrated that this putative short C9ORF72 isoform does not interact with SMCR8 or WDR41 (Fig EV1B). Co-immunoprecipitation experiments also indicated that C9ORF72 and SMCR8 bind to HSC70, but not to BAG3 (Fig EV1C), suggesting that BAG3 recruitment is not direct but mediated by HSC70. Whether C9ORF72 is a client or a regulator of the HSC70 and BAG3 pathway remains to be determined.

C9ORF72 in complex with SMCR8 and WDR41 interacts with RAB8a and RAB39b

Proteomic analysis revealed association of C9ORF72 with Rab8a and Rab39b (Fig EV1A). However, C9ORF72 was reported to interact with Rab1, Rab5, Rab7, and Rab11 (Farg *et al*, 2014). Since our proteomic analysis was performed using Neuro-2A cell lysate that may incompletely represent the proteomic complexity of a tissue, we tested the association of C9ORF72 with various other Rab GTPases. Among the Rab proteins tested, we confirmed that the C9ORF72 complex interacts with RAB8a and weakly with RAB39b, as well as, but in a lesser extent, with their paralogs RAB8b and RAB39a (Fig 1D). Of interest, RAB8 is involved in a *Drosophila* model of FTD (West *et al*, 2015) and interacts with OPTN and TBK1 that are mutated in ALS (Maruyama *et al*, 2010; Cirulli *et al*, 2015; Freischmidt *et al*, 2015). Also, mutations in the X-linked *RAB39b* gene lead to intellectual disability associated with autism, epilepsy, and early-onset parkinsonism (Giannandrea *et al*, 2010; Wilson *et al*, 2014). Among other Rab tested, the C9ORF72 complex presents also some weak interaction with RAB6a, RAB12, RAB25, RAB33a, and RAB38 (Fig 1D). In contrast, C9ORF72 alone or in complex with SMCR8 present no or very little interaction with RAB1, RAB5, RAB7, and RAB11 (Fig 1D). Similarly, we found no interaction of RAGA/D or RALB GTPases with C9ORF72 alone or in complex with SMCR8. Of technical interest, co-transfection of WDR41 abolished expression of various Rab cDNAs; thus, co-immunoprecipitations were performed with the C9ORF72 and SMCR8 proteins (Fig 1D). Interactions between RAB8a or RAB39b and the C9ORF72, SMCR8, and WDR41 complex were not inhibited by mutations locking Rab proteins in a constitutive GDP-bound inactive or a constitutively GTP-bound active form. Further co-immunoprecipitation assays indicated that the interaction between C9ORF72 complex and RAB8a or RAB39b is mainly dependent on the presence of SMCR8 since C9ORF72 or WDR41 alone immunoprecipitated little RAB8a or no RAB39b (Fig EV1D). In that aspect, SMCR8 expressed alone immunoprecipitated RAB8a and RAB39b but also some other Rab GTPases, including RAB24, RAB32, and RAB7L1, which is also known as RAB29 (Fig EV1E). These interactions were lost when SMCR8 was in complex with C9ORF72, suggesting that the specificity of SMCR8 toward Rab GTPase proteins may change according to the SMCR8 partners.

Since these experiments were performed in transfected cells, we next tested whether the endogenous C9ORF72 complex may immunoprecipitate endogenous RAB8a and RAB39b. Taking advantage of the recombinant HIS-tagged C9ORF72, SMCR8, and WDR41 complex purified from baculovirus-infected insect cells, we immunized mice, but failed to obtain antibodies specific to C9ORF72

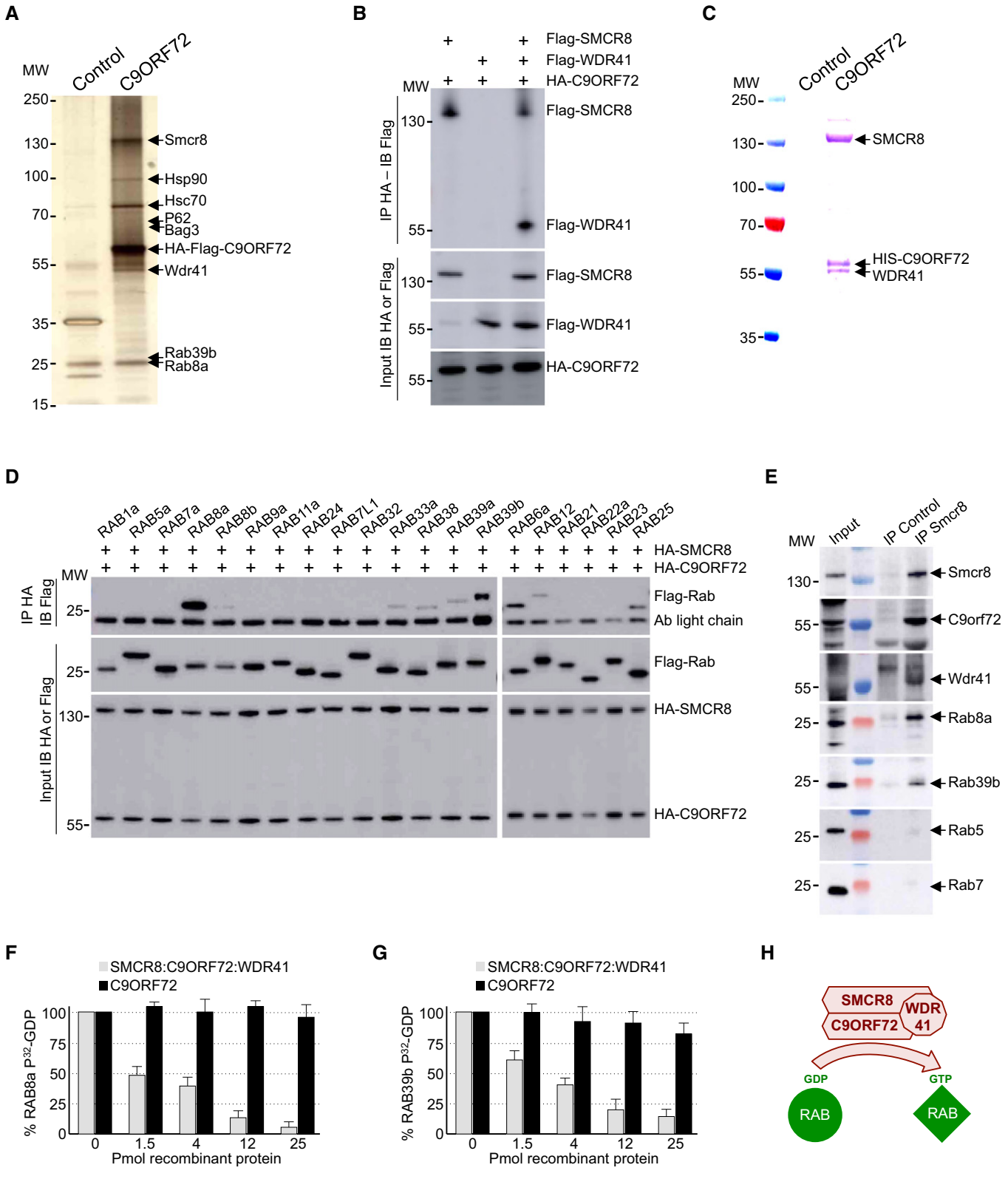


Figure 1.

despite many attempts. In contrast, we successfully developed a monoclonal antibody (1D2) against SMCR8. Immunoblotting indicated that Smcr8, Rab8a, and Rab39b are all expressed in mouse brain as well as in cultures of primary E18 cortical mouse neurons (Fig EV1F). We were not able to test expression of endogenous C9orf72 and Wdr41 due to the poor quality of the commercial

antibodies tested, but expression of C9orf72 in mouse brain had been reported previously (Suzuki *et al*, 2013). Importantly, immunoprecipitation of endogenous Smcr8 from total mouse brain lysate successfully pulled down Smcr8, but also endogenous C9orf72 and Wdr41, as well as Rab8a and Rab39b (Fig 1E). In contrast, we did not detect endogenous Rab5 or Rab7 in Smcr8 immunoprecipitation (Fig 1E), thus confirming our transfection experiments.

C9ORF72 in complex with SMCR8 and WDR41 is a GEF for Rab GTPase

Both C9ORF72 and SMCR8 contain DENN domains characteristic of Rab-guanine nucleotide exchange factors (Zhang *et al*, 2012; Levine *et al*, 2013). Thus, we tested whether the complex composed of C9ORF72, SMCR8, and WDR41 presents any GEF activity. Purified recombinant GST-tagged RAB8a was preloaded with ^{32}P-labeled GDP, and nucleotide release was monitored in the presence of increasing amount of C9ORF72 recombinant complex purified from baculovirus-infected insect cells. The complex formed by C9ORF72, SMCR8, and WDR41 stimulated GDP release since more than 90% of 40 pmol of RAB8a released its associated GDP in the presence of 25 pmol of C9ORF72, SMCR8, and WDR41 complex (Fig 1F). In contrast, addition of recombinant purified C9ORF72 alone had no or very little effect, suggesting that C9ORF72 is active only when in complex with SMCR8 (Fig 1F). As the C9ORF72 complex also interacts with RAB39b, we tested its GEF activity and found that, similar to RAB8a, the complex composed of C9ORF72, SMCR8, and WDR41 activated GDP release from recombinant purified GST-RAB39b (Fig 1G). In contrast, addition of recombinant purified C9ORF72 alone had no or little effect (Fig 1G). As control, the complex formed by C9ORF72, SMCR8, and WDR41, which does not interact with RAB32 or RAB29, presented no or very little GEF activity toward these Rab GTPases (Fig EV1G). Overall, these results demonstrate that C9ORF72 forms a specific complex with SMCR8 and WDR41, which acts as a GEF effector for at least RAB8a and RAB39b (Fig 1H).

RAB8a and RAB39b interact with the autophagy receptors P62

Proteomic analysis identified that C9ORF72 potentially interacts with P62, which is encoded by the *SQSTM1* gene and targets poly-ubiquitinated proteins for autophagy degradation (Table EV1). Co-immunoprecipitation experiments confirmed that the complex formed by C9ORF72, SMCR8, and WDR41 interacts weakly with P62, but also with OPTN that is alike P62, an autophagy receptor. Further co-immunoprecipitation assays pinpointed that among the proteins of the C9ORF72 complex, P62 mainly interacts with WDR41 and SMCR8 (Fig EV2A), while OPTN mainly interacts with WDR41 (Fig EV2B). Of interest, mutations in *SQSTM1* or *OPTN* cause ALS (Maruyama *et al*, 2010; Fecto *et al*, 2011), and *OPTN* is a potential modifier gene for FTD (Pottier *et al*, 2015). We noted that the quantities of P62 or OPTN immunoprecipitated by the C9ORF72 complex were low, suggesting a potential indirect association. Since OPTN interacts with RAB8 (Hattula & Peränen, 2000; Pilli *et al*, 2012), we tested whether interaction of the C9ORF72 complex with P62 would be mediated by intermediate proteins, namely RAB8 or RAB39. Indeed, both HA-tagged RAB8a and RAB39b readily immunoprecipitated Flag-tagged P62 (Fig EV2C).

Decreased expression of C9ORF72 alters autophagy

C9ORF72 is interacting with various proteins connected to autophagy (Table EV1; Behrends *et al*, 2010; Farg *et al*, 2014). Thus, we tested whether depletion of *C9ORF72* expression modifies autophagy in neuronal cells. First, we evaluated the formation of vesicles containing LC3B, a protein encoded by the *MAP1LC3B* gene, which is specifically lipidated and localized to autophagic vesicles. Expression of double-tagged GFP-RFP-LC3B in primary cultures of embryonic mouse cortical neurons shows diffuse cytoplasmic localization and rare punctuate structures. In contrast, treatment with torin, which inhibits the autophagic-inhibitory mTOR pathway, induces formation of LC3B puncta (Fig 2A). Importantly, shRNA-mediated depletion of *C9orf72* abolishes this activation of autophagy (Fig 2A). GFP-RFP-LC3B expression in primary neuronal cultures revealed no differences in the GFP/RFP ratio. This ratio reflects the rate of formation of autophagosome relative to their fusion to lysosome, suggesting that C9ORF72 acts on the formation of autophagosome rather than on their degradation by the lysosome. Similar inhibitory effect of *C9orf72* siRNA on autophagy activation was observed with primary neuronal cultures treated with rapamycin instead of torin, or when GT1-7 cells, a transformed murine neuronal cell line expressing similar level of *C9orf72* compared to primary neuronal cell cultures, were tested (Fig EV2D). Next, we confirmed these results by investigating the ratio of lipidated LC3B. Western blotting analysis revealed that the level of phosphatidylethanolamine-linked LC3B-II was slightly decreased in primary cultures of embryonic mouse cortical neurons transduced with lentivirus expressing shRNAs against *C9orf72* in basal condition (Fig 2B). This inhibitory

Figure 2. Reduced expression of C9ORF72 partly impairs autophagy.

A Left panel, representative images of organotypic cultures of E18 mouse cortical neurons transfected with GFP-RFP-LC3B and transduced with lentivirus expressing either control shRNA or shRNA targeting *C9orf72* mRNA and treated or not with Torin. Right panel, quantification of LC3B puncta.

B Upper panel, immunoblot analysis of endogenous LC3B (Map1lc3b), C9orf72, and control actin of E18 mouse cortical neurons transduced with lentivirus expressing either control shRNA or shRNA targeting *C9orf72* mRNA and treated or not with Torin. Lower panel, real-time RT–qPCR quantification of endogenous *C9orf72* mRNA expression relative to *Rplp0* mRNA.

C Left panel, representative images of immunofluorescence labeling of endogenous P62 (Sqstm1) on organotypic cultures of E18 mouse cortical neurons transduced with lentiviral particles expressing either control shRNA or shRNA targeting *C9orf72*. Right panel, quantification of P62 aggregates.

D Left panel, representative images of immunofluorescence labeling of transfected constructs (green) and endogenous P62 (Sqstm1, red) on GT1-7 neuronal cells transfected with siRNA targeting *C9orf72* and plasmids expressing control GFP, and HA-tagged wild-type or mutant RAB39b (CA, Q68L or CN, S22N), RAB8a, or RAB7. Right panel, quantification of P62 aggregates.

Data information: Scale bars, 10 μm. Nuclei were counterstained with DAPI. Error bars indicate SEM. Student's *t*-test, $*P < 0.05$, $***P < 0.001$, $n = 3$.
Source data are available online for this figure.

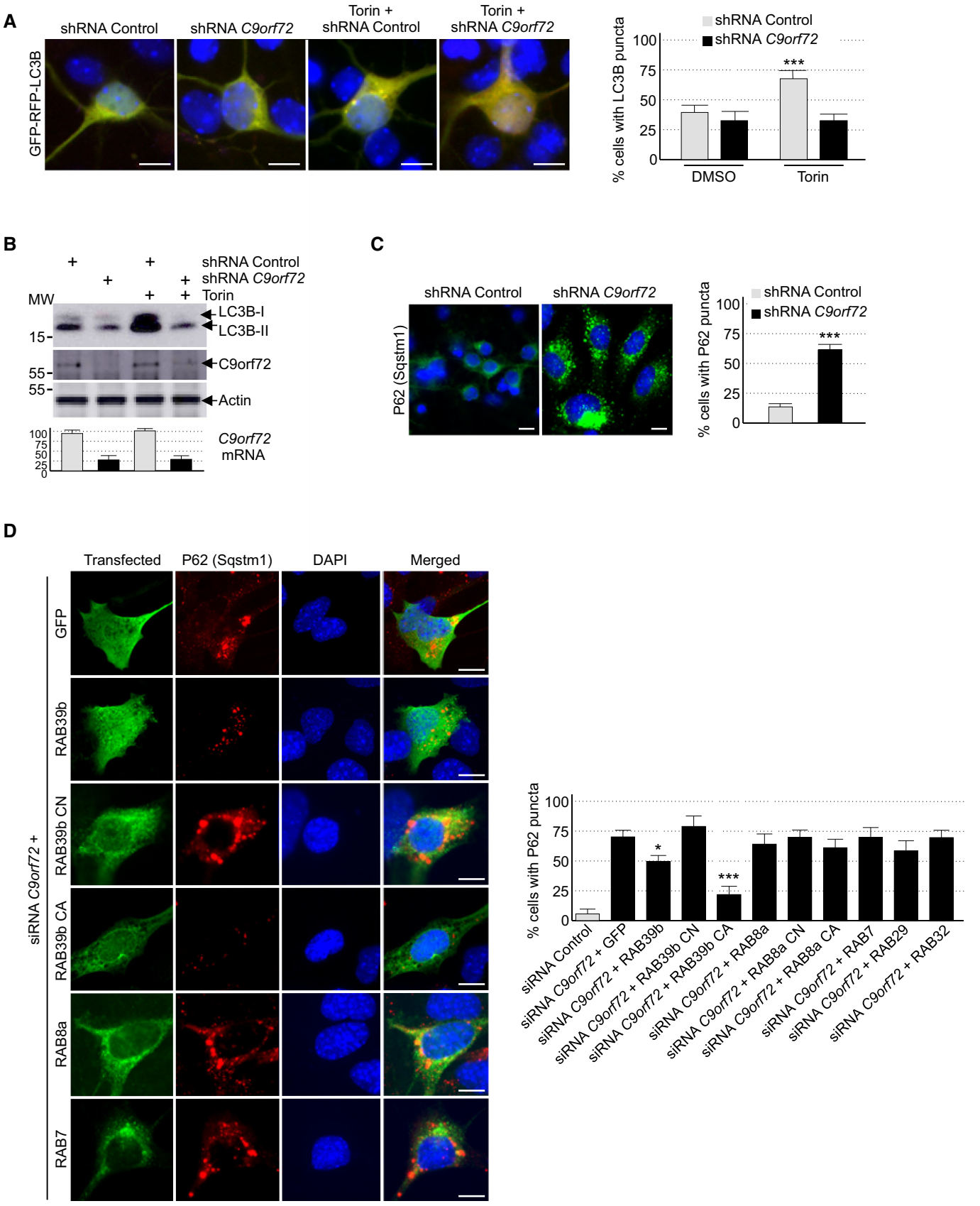


Figure 2.

effect was much more evident in torin-treated neurons since activation of the lipidation of LC3B was abolished by shRNA-mediated depletion of C9orf72 (Fig 2B). Identical results were observed in neurons treated with rapamycin instead of torin, or in GT1-7 neuronal cells (Fig EV2E). Depletion of at least 50% of C9orf72 expression was confirmed by Western blotting and, due to the poor quality of the commercial antibodies used, further confirmed by RT–qPCR (Figs 2B and EV2E). Also, we found that decreased expression of C9orf72 reduced but not totally abolished accumulation of LC3B-II in neuronal cells treated with bafilomycin A1, an inhibitor of autophagosome degradation (Fig EV2E). Overall, these data suggest that depletion of C9ORF72 has a partial deleterious effect on autophagy.

Autophagy is a crucial mechanism to clear misfolded or aggregated proteins. Thus, we tested whether depletion of C9ORF72 had any effect on the degradation of protein aggregate. The P62 protein, which is encoded by the *SQSTM1* gene, bridges aggregates of poly-ubiquitinated proteins to LC3B, thereby targeting these proteins toward autophagy. Furthermore, P62-positive aggregates are a histological hallmark of ALS-FTD patients with expansion of GGGGCC repeats in *C9ORF72* (Al-Sarraj *et al*, 2011). Immunofluorescence labeling of P62 indicated an accumulation of unresolved P62 aggregates upon depletion of *C9orf72* in primary cultures of embryonic mouse cortical neurons (Fig 2C). Identical results were obtained for GT1-7 neuronal cells or when partners of C9orf72, namely Smcr8 and Wdr41, were siRNA-depleted (Fig EV2F). As control, expression of optimized HA-tagged C9ORF72, which has a nucleotide sequence resistant to the siRNAs targeting endogenous mouse *C9orf72* mRNA, fully rescued autophagy dysfunction caused by siRNA-mediated depletion of C9orf72 (Fig EV2G). Of interest, while the long isoform of C9ORF72 rescued autophagy dysfunction, the short form of C9ORF72 was inactive in that respect (Fig EV2G). These results suggest that the short isoform of C9ORF72 is either a null variant or possesses a cellular function unrelated to autophagy.

RAB39b corrects autophagy dysfunction caused by depletion of C9ORF72

Depletion of C9ORF72 alters autophagy, and C9ORF72 in complex with SMCR8 interacts and promotes GDP/GTP exchange of RAB8a and RAB39b, which are two Rab GTPases involved in autophagy (Pilli *et al*, 2012; Seto *et al*, 2013). This questions the relevance of the interaction of C9ORF72 with RAB8a or RAB39b, notably to control autophagy. Importantly, expression of a constitutively active (Q68L, mutant CA) form of RAB39b, which is locked in its GTP conformation and consequently does not require any GEF activity, fully corrected autophagy alteration caused by a decreased expression of C9orf72 (Fig 2D). Expression of wild-type HA-tagged RAB39b had only a partial effect on autophagy rescue, while a GDP-locked constitutively negative (S22N, mutant CN) RAB39b did not rescue autophagy alterations caused by siRNA-mediated depletion of C9orf72. As control of RAB39b specificity, expression of RAB7, RAB29, RAB32, or wild-type, constitutively active or inactive RAB8A did not rescue autophagy alteration caused by C9orf72 depletion (Fig 2D). These results suggest that the GEF activity of the C9ORF72 complex toward RAB39b is important to control autophagy in neuronal cells.

TBK1 phosphorylates SMCR8

Tandem-tag purification of C9ORF72 identified a putative weak interaction with TANK-Binding Kinase 1 (TBK1) (Table EV1). Of interest, mutations in *TBK1* lead to ALS (Cirulli *et al*, 2015; Freischmidt *et al*, 2015), and RAB8, which interacts with the C9ORF72 complex, is involved in the autophagic elimination of intracellular pathogens through the kinase TBK1 (Pilli *et al*, 2012). Thus, we tested whether the C9ORF72 complex would interact with the TBK1 kinase complex. Co-immunoprecipitation experiments indicated that the complex formed by C9ORF72 and SMCR8 did not bind to TBK1 directly, but interacted with all three TBK1 adaptor proteins (Fig 3A), namely TANK, SINTBAD (*TBKBP1*), and NAP1 (*AZI2*). These adaptor proteins are essential to direct TBK1 toward specific cellular comportments and functions, notably autophagy (Goncalves *et al*, 2011). These interactions question whether TBK1 phosphorylates C9ORF72. *In vitro* kinase assay demonstrated that TBK1 phosphorylates SMCR8, but not C9ORF72 or WDR41 (Fig 3B). We repeated that experiment using purified TBK1 overexpressed from HEK293 cells and recombinant HIS-C9ORF72, SMCR8, and WDR41 complex purified from baculovirus-infected insect cells. Mass spectrometry analysis identified SMCR8 serine 402 and threonine 796 as TBK1 phosphorylation sites (Fig 3C). Consistent with the consensus motif identified in other substrates of TBK1 (Ma *et al*, 2012), both SMCR8 serine 402 and threonine 796 are followed by a leucine.

Figure 3. SMCR8 is phosphorylated by TBK1.

A Immunoblot analysis of HA-immunoprecipitated proteins and lysate of HEK293 cells co-expressing HA-tagged C9ORF72 and HA-tagged SMCR8 with Flag-tagged TBK1, NAP1, TANK, or SINTBAD.

B Immunoprecipitated HA-tagged C9ORF72, HA-tagged SMCR8, and HA-tagged WDR41 expressed in HEK293 were subjected to *in vitro* TBK1 kinase assay in the presence of γ-^{32}P-radiolabelled ATP. Proteins were separated by SDS–PAGE migration and phosphorylation was detected by autoradiography (upper panel), while expression was detected by Western blotting (lower panel).

C Mass spectrometry identification of SMCR8 phosphorylation sites by *in vitro* TBK1 kinase assay of HIS-tagged C9ORF72:SMCR8:WDR41 complex purified from baculovirus-infected insect cells.

D Upper panel, representative images of immunofluorescence labeling of transfected constructs (green) and endogenous P62 (Sqstm1, red) on GT1-7 neuronal cells transfected with siRNA targeting the 3′UTR of *Smcr8* mRNA and plasmids expressing control GFP and HA-tagged wild-type or mutant SMCR8 (TA, S402A and T796A; TD, S402D and T796D; UD, S400D, S492D, S562D, and T666D). Lower panel, quantification of P62 aggregates.

E Upper panel, representative images of immunofluorescence labeling of transfected constructs (green) and endogenous P62 (Sqstm1, red) on GT1-7 neuronal cells transfected with siRNA targeting *Tbk1* and plasmids expressing control GFP and HA-tagged wild-type or mutant SMCR8 or RAB39b. Lower panel, quantification of P62 aggregates.

Data information: Scale bars, 10 μm. Nuclei were counterstained with DAPI. Error bars indicate SEM. Student's *t*-test, $*P < 0.05$, $***P < 0.001$, $n = 3$.
Source data are available online for this figure.

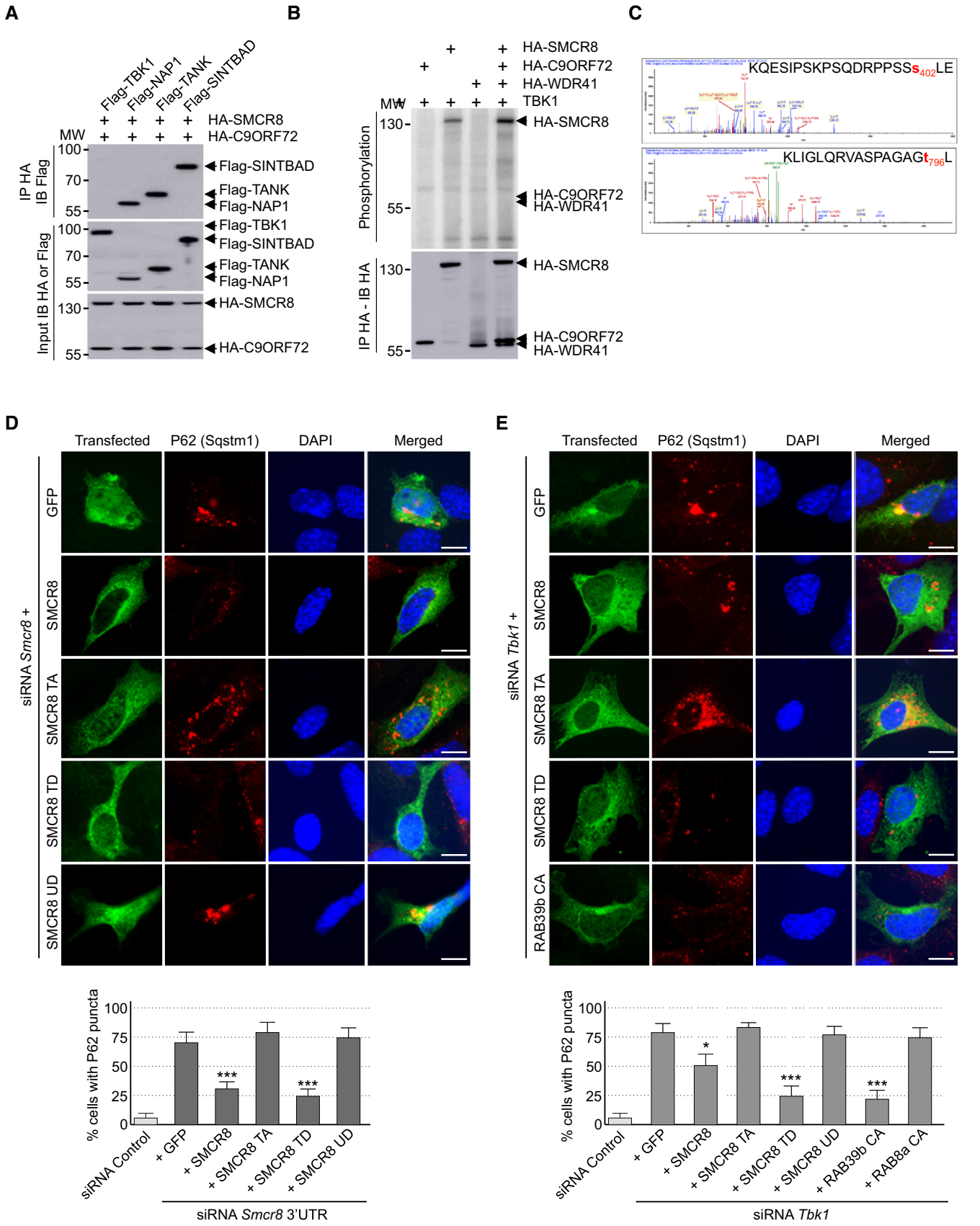

A
Flag-TBK1
Flag-NAP1
Flag-TANK
Flag-SINTBAD
HA-SMCR8
HA-C9ORF72
MW
IP HA IB Flag
Flag-SINTBAD
Flag-TANK
Flag-NAP1
Input IB HA or Flag
Flag-TBK1
HA-SMCR8
HA-C9ORF72
B
HA-SMCR8
HA-C9ORF72
HA-WDR41
TBK1
Phosphorylation
IP HA - IB HA
C
KQESIPSKPSQDRPPSSs402LE
KLIGLQRVASPAGAGt796L
D
Transfected
P62 (Sqstm1)
DAPI
Merged
siRNA Smcr8 +
GFP
SMCR8
SMCR8 TA
SMCR8 TD
SMCR8 UD
% cells with P62 puncta
siRNA Control
+ GFP
+ SMCR8
+ SMCR8 TA
+ SMCR8 TD
+ SMCR8 UD
siRNA Smcr8 3'UTR
E
siRNA Tbk1 +
RAB39b CA
+ RAB39b CA
+ RAB8a CA
siRNA Tbk1

Figure 3.

TBK1 phosphorylation of SMCR8 is important for autophagy

We constructed TBK1 phospho-dead (S402A, T796A) and phosphomimetic (S402D, T796D) mutants of SMCR8 and noted that these mutants were normally localized and expressed in transfected neuronal cells. Also, phospho-dead and phosphomimetic mutants of SMCR8 immunoprecipitated C9ORF72 and WDR41 as well as wild-type SMCR8, suggesting that mutation of SMCR8 S402 and T796 did not alter expression, stability, structure, and ability of SMCR8 to form a complex. Next, we tested whether SMCR8 phosphorylation was important for its cellular function. Depletion of Smcr8 altered autophagy as illustrated by accumulation of P62 aggregates in neuronal GT1-7 cells transfected with a siRNA targeting the 3′UTR of *Smcr8* mRNA (Fig 3D). Depletion of endogenous Smcr8 expression was confirmed by immunoblotting (Fig EV3A). As control, co-transfection of a siRNA-resistant HA-tagged SMCR8 cDNA corrected autophagy dysfunction due to reduced expression of Smcr8 (Fig 3D). Importantly, expression of HA-tagged S402D and T796D double mutant of SMCR8 (mutant TD), which mimics a constitutive TBK1 phosphorylation of SMCR8, also corrected autophagy dysfunction caused by siRNA-mediated depletion of Smcr8 (Fig 3D). In contrast, a phospho-dead (S402A, T796A; mutant TA) SMCR8 was unable to rescue autophagy alteration (Fig 3D), demonstrating importance of SMCR8 phosphorylation for its function. As a further control, a phosphomimetic mutant of SMCR8 unrelated to TBK1 (SMCR8 S400D, S492D, S562D, T666D; mutant UD) did not correct autophagy alteration caused by siRNA-mediated depletion of Smcr8 (Fig 3D).

We then investigated whether SMCR8 phosphorylation was important for TBK1 cellular function. siRNA-mediated depletion of Tbk1 promoted accumulation of P62 aggregates in neuronal GT1-7 cells (Fig 3E). Decreased expression of Tbk1 was confirmed by immunoblotting (Fig EV3B). Of interest, expression of wild-type HA-tagged SMCR8 rescued only partially the dysfunction of autophagy caused by Tbk1 depletion (Fig 3E). In contrast, transfection of the S402D and T796D mutants of SMCR8 (mutant TD), which mimics a constitutive TBK1 phosphorylation, fully corrected autophagy alteration caused by Tbk1 depletion (Fig 3E). As controls, a phospho-dead (S402A, T796A; mutant TA) as well as a phosphomimetic mutant of SMCR8 unrelated to TBK1 (SMCR8 S400D, S492D, S562D, T666D; mutant UD) was unable to rescue autophagy alteration caused by siRNA-mediated depletion of Tbk1 (Fig 3E). Importantly, depletion of Tbk1 was also corrected by expression of constitutively active GTP-locked HA-tagged RAB39b, but not by constitutively active RAB8A (Fig 3E). Overall, these results suggest that phosphorylation of SMCR8 by TBK1 is important to control autophagy, but also that TBK1, C9ORF72 complex, and RAB39b belong to a common pathway regulating autophagy in neuronal cells.

ULK1 phosphorylates SMCR8

We noted that SMCR8 was initially identified as an interactant of the ULK1 kinase complex, which initiates autophagy (Behrends *et al*, 2010). Similarly, our tandem-tag purification of C9ORF72 identified a potential weak interaction with a component of the ULK1 kinase complex, namely Fip200, encoded by the *Rb1cc1* gene (Table EV1). Co-immunoprecipitation experiments confirmed that the complex formed by C9ORF72, WDR41, and SMCR8 interacts weakly with ULK1 and that this interaction is mostly dependent on the presence of SMCR8 (Fig EV3C). Similar to TBK1, *in vitro* kinase assay demonstrated that ULK1 phosphorylates SMCR8 alone or in complex, but not C9ORF72 or WDR41 (Fig EV3D). Mass spectrometry analysis identified SMCR8 serine 400, serine 492, serine 562, and threonine 666 or serine 667 as ULK1 phosphorylation sites (Fig EV3E). However, depletion of Ulk1 or of both Ulk1 and Ulk2 using siRNA in neuronal GT1-7 cells leads to only partial accumulation of P62 aggregates compared to siRNA-mediated depletion of Tbk1 (Fig EV3F). Furthermore, a phosphomimetic mutant of SMCR8 (S400D, S492D, S562D, T666D; mutant UD) simulating constitutive phosphorylation of SMCR8 by the ULK1 kinase did not correct autophagy alteration caused by depletion of Smcr8 (Fig 3D). Overall, these results suggest that in GT1-7 neuronal cells and in the time frame of our study, phosphorylation of SMCR8 by the TBK1 kinase plays a crucial role to regulate autophagy compared to ULK1.

Decreased expression of C9ORF72 promotes aggregation of TDP-43

Brain sections of ALS-FTD patients are characterized by the presence of neuronal cytoplasmic inclusions containing the TAR DNA-binding protein 43 (TDP-43) that is abnormally ubiquitinated, phosphorylated, and truncated (Arai *et al*, 2006; Neumann *et al*, 2006). Since aggregates of TDP-43 are resolved by autophagy (Filimonenko *et al*, 2007; Ju *et al*, 2009; Urushitani *et al*, 2010; Wang *et al*, 2012; Barmada *et al*, 2014; Scotter *et al*, 2014), we tested whether decreased expression of *C9orf72* had any effect on TDP-43. Importantly, Western blotting assays demonstrated that shRNA-mediated depletion of *C9orf72* in mouse embryonic cortical neurons induced the accumulation of a ~30 kDa truncated fragment of Tdp-43 (Fig 4A). Similarly, immunofluorescence analysis against phosphorylated serine 409/410 of Tdp-43 confirmed accumulation of cytoplasmic aggregates of Tdp-43 in neurons depleted of C9orf72 (Fig 4B). Note that the accumulation of Tdp-43 aggregates was evident in primary neuronal cultures upon 7 days of transduction with C9orf72 shRNA lentivirus, while we did not detect accumulation of Tdp-43 aggregates in GT1-7 neuronal cells treated with C9orf72 siRNA for 24 h, suggesting that accumulation of Tdp-43 may be cell or time dependent. Finally, mutations in *TARDBP*, encoding TDP-43, lead to aggregate of TDP-43 proteins and cause ALS-FTD (Gitcho *et al*, 2008; Kabashi *et al*, 2008; Rutherford *et al*, 2008; Sreedharan *et al*, 2008; Van Deerlin *et al*, 2008; Yokoseki *et al*, 2008). We found that overexpression of C9ORF72 reduced the aggregation of the D169G mutant of TDP-43 (Fig 4C). These results are consistent with previous works showing that enhancing autophagy corrects aggregation and toxicity of mutant TDP-43 (Wang *et al*, 2012; Barmada *et al*, 2014).

Phosphomimetic mutants of SMCR8 and constitutively active GTP-locked RAB39b correct autophagy alteration due to depletion of Smcr8, C9orf72, or Tbk1, suggesting that these proteins belong to the same pathway. Since TDP-43 regulates indirectly autophagy (Bose *et al*, 2011; Xia *et al*, 2016), we tested whether C9ORF72 or RAB39b can rescue autophagy misregulation caused by reduced expression of *Tardbp*. Depletion of Tdp-43 partly altered autophagy as illustrated by accumulation of P62 aggregates in neuronal GT1-7 cells transfected with a siRNA targeting *Tardbp* (Fig EV4). However, we found no correction of autophagy upon transfection of C9ORF72, SMCR8, or wild-type or constitutively active RAB8A or RAB39b

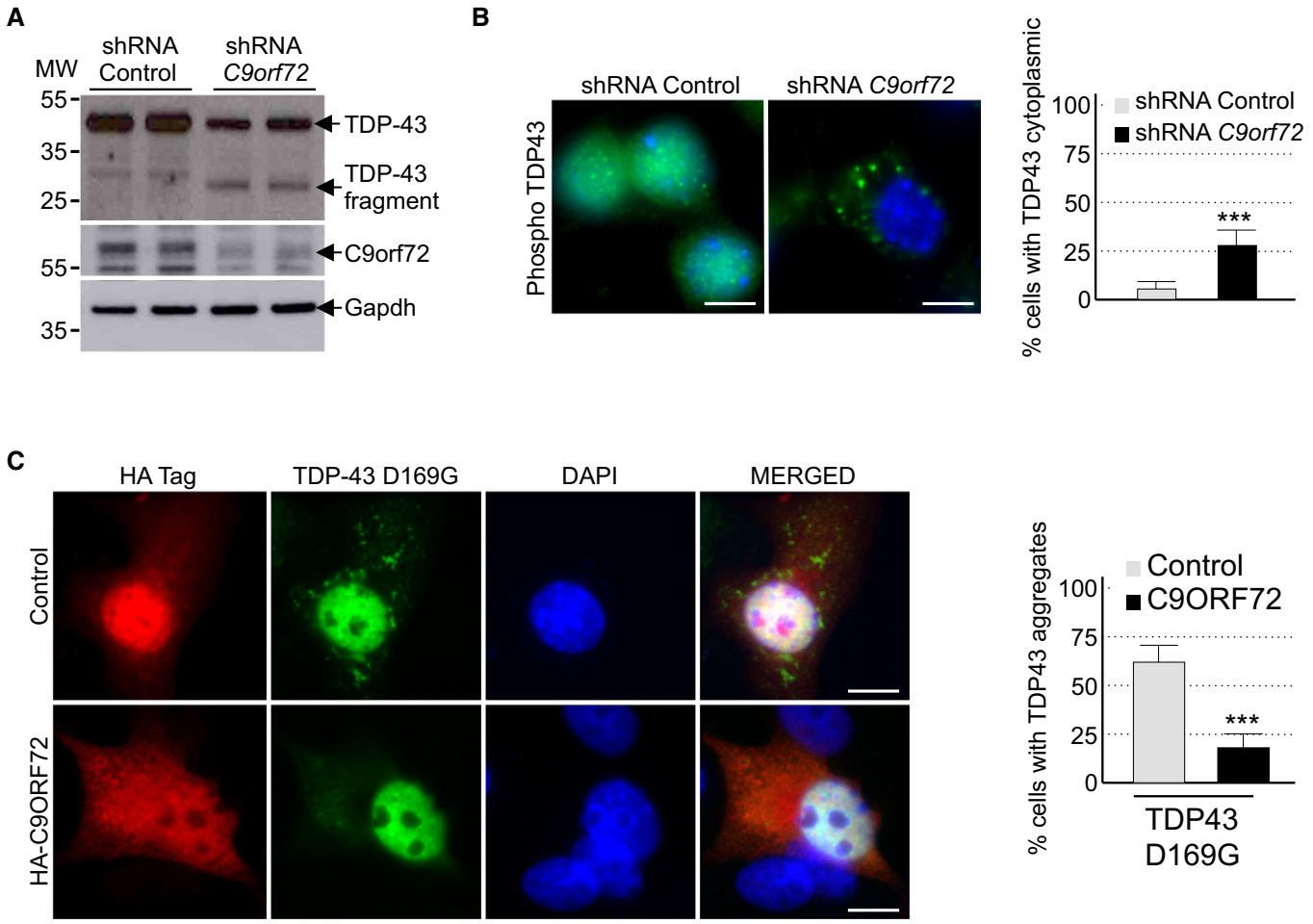


Figure 4. Reduced expression of C9ORF72 promotes aggregation of TDP-43.

A Immunoblot analysis of endogenous Tdp-43, C9orf72, and control Gapdh of E18 mouse cortical neurons transduced with lentivirus expressing either control shRNA or shRNA targeting *C9orf72* mRNA.

B Left panel, representative images of immunofluorescence labeling of endogenous phosphorylated Ser409/410-Tdp-43 on primary cultures of E18 mouse cortical neurons transduced with lentiviral particles expressing either control shRNA or shRNA targeting *C9orf72*. Right panel, quantification of cytoplasmic Tdp-43 aggregates.

C Left panel, representative images of immunofluorescence labeling of D169G mutant GFP-tagged TDP-43 on primary cultures of E18 mouse cortical neurons transfected with either HA-tagged control or HA-C9ORF72 plasmid. Right panel, quantification of cells with cytoplasmic aggregates of TDP-43.

Data information: Scale bars, 10 μm. Nuclei were counterstained with DAPI. Error bars indicate SEM. Student's *t*-test, $***P < 0.001$, $n = 3$.
Source data are available online for this figure.

(Fig EV4). These negative results highlight the specificity of RAB39b toward TBK1 and C9ORF72, but also indicate that while TDP-43 regulates autophagy, this regulation is independent or downstream of the TBK1, C9ORF72, and RAB39b pathway. This is consistent with the recent report that loss of TDP-43 downregulates Dynactin 1 mRNA, which impairs autophagy at the late step of fusion of autophagosomes to lysosomes (Xia *et al*, 2016).

Decreased expression of C9ORF72 synergizes Ataxin-2 Q30x toxicity

Since decreased expression of C9ORF72 inhibits autophagy and promotes accumulation of aggregates of TDP-43, we searched for other proteins prone to aggregation that may accumulate upon depletion of C9ORF72. Ataxin-2 (encoded by the gene *ATXN2*) is a cytoplasmic RNA-binding protein that interacts with the poly(A)-binding protein (PABP) and regulates mRNA stability (Kozlov *et al*, 2001; Yokoshi *et al*, 2014). Abnormal expansion over 34 glutamines in *ATXN2* leads to spinocerebellar ataxia type 2 (SCA2) (Imbert *et al*, 1996; Pulst *et al*, 1996; Sanpei *et al*, 1996), while intermediate expansion of polyQ (27–33 repeats) is an increased risk of ALS-FTD (Elden *et al*, 2010; Daoud *et al*, 2011; Ross *et al*, 2011; Van Damme *et al*, 2011; Lattante *et al*, 2014). We found that both HA-tagged Ataxin-2 with control length (Q22x) and intermediate size (Q30x) of polyQ localized diffusely into the cytoplasm of primary cultures of embryonic mouse cortical neurons (Fig 5A). Importantly, shRNA-mediated decreased expression of C9orf72 promoted accumulation of aggregates of Ataxin-2 with intermediate polyQ length in primary cultures of neurons (Fig 5A). Similar results were observed in GT1-7 neuronal cells or when

partners of C9orf72, namely Smcr8 and Wdr41, were siRNA-depleted (Fig EV5A). As a control, treatment of neuronal cells with bafilomycin A1, a drug that blocks autophagy, also promoted accumulation of Ataxin-2 Q30x aggregates (Fig EV5A). In contrast, depletion of C9orf72, Smcr8, or Wdr41 or bafilomycin treatment had no effect on the diffuse localization of Ataxin-2 with control (Q22x) polyQ size (Figs 5A and EV5A). Immunoblotting confirmed identical expression of HA-tagged Ataxin-2 Q22x upon siRNA-mediated depletion of C9orf72, Smcr8, or Wdr41 or bafilomycin treatment compared to control siRNA (Fig EV5B). Similarly, expression of HA-tagged Ataxin-2 Q30x is identical between control and siRNA-treated C9orf72, Smcr8, or Wdr41, but slightly increased upon bafilomycin A1 treatment (Fig EV5B). Consistent with a previous observation of an increased stability of Ataxin-2 with intermediate size of polyQ (Elden *et al*, 2010), we noted that HA-tagged Ataxin-2 Q30x is expressed at twofold to threefold higher levels than HA-tagged Ataxin-2 Q22x (Fig EV5B and C). As further controls, the effect of C9orf72 loss on aggregation of Ataxin-2 Q30x was specific since we did not observe increased aggregation of mutant SOD1, mutant FUS, huntingtin, or Ataxin-3 with polyQ expansion upon siRNA-mediated depletion of *C9orf72* expression, at least in transfected cells and in the time frame of our study (Fig EV5D).

Next, we investigated whether decrease expression of C9ORF72 leads to any reduction in neuronal viability. Consistent with a previous report (Wen *et al*, 2014), neither siRNA- nor shRNA-mediated decreased expression of *C9orf72* in neuronal primary cultures induced significant cell death, at least in the conditions and time frame of our study (Fig 5B). Similarly, expression of Ataxin-2 with control (Q22x) or intermediate (Q30x) size of polyQ had little effect on neuronal viability (Fig 5B). In contrast, decreased expression of C9orf72 and simultaneous expression of Ataxin-2 with intermediate size of polyQ (Q30x) induced neuronal cell death (Fig 5B). As a control, siRNA-mediated depletion of C9orf72 with concomitant expression of control Ataxin-2 (Q22x) did not reduce neuronal cell viability (Fig 5B).

To confirm these results *in vivo*, we developed zebrafish models with decreased expression of C9orf72 and which express Ataxin-2 with either normal or intermediate size of polyQ. Of technical interest, we used a reduced quantity (50%) of antisense morpholino oligonucleotides (AMOs) known to block translation of the zebrafish ortholog of C9ORF72 compared to a previous study (Ciura *et al*, 2013). In these conditions, we observed little toxicity and no abnormal motor phenotype associated with a reduction in C9orf72 (Fig 5C). Quantification of endogenous zebrafish *C9orf72* mRNA by RT–qPCR confirmed a partial (50–60%) decreased expression of C9orf72 upon antisense AMO injection compared to control conditions (Fig EV5E). We also controlled by RT–qPCR the equal expression of HA-tagged Ataxin-2 Q22x or Q30x in injected zebrafish (Fig EV5F). Consistent with previous results (Elden *et al*, 2010), we noted a higher expression of Ataxin-2 Q30x compared to Ataxin-2 Q22x (Fig EV5F). Importantly, we observed no toxicity associated with the sole expression of Ataxin-2 with normal (Q22x) or intermediate (Q30x) length of polyQ (Fig EV5G). In contrast, the decreased expression of C9orf72 associated with the expression of Ataxin-2 with intermediate size of polyQ (Q30x) resulted in an abnormal motor behavior, specifically a reduced touch-evoked escape response (Fig 5C). Indeed, after stimulation by light touch to the fish tail, swimming distance, average velocity, and maximum velocity were significantly reduced in zebrafish embryos knocked down for C9orf72 and expressing Ataxin-2 with thirty glutamines (Fig 5D–F). As a control, decreased expression of C9orf72 with concomitant expression of Ataxin-2 with control length of polyQ (Q22x) had no significant pathogenic effect on swimming episodes following escape response (Fig 5D–F). Furthermore, the specificity of the phenotype was confirmed by injection of the same concentration of a control mismatch morpholino, which did not cause any locomotor phenotype (Fig EV5G–J). As further controls, expression of Ataxin-2 with either control or intermediate length of polyQ alone or with a mismatch antisense morpholino oligonucleotide also did not cause any locomotor phenotypes (Fig EV5G–J).

Finally, we analyzed the morphology of the axonal projections from spinal motor neurons by immunofluorescence against the synaptic vesicle marker Sv2. Similar to results presented above, knockdown of C9orf72 alone or the sole expression of Ataxin-2 (Q22x or Q30x) had no effect compared to control non-injected or mismatch AMO-injected fish (Fig 5G). Also, knockdown of C9orf72 with concomitant expression of Ataxin-2 with control size (Q22x) of polyQ had no toxic effect (Fig 5G and H). In contrast, the decreased expression of C9orf72 with simultaneous expression of Ataxin-2 with intermediate size of polyQ (Q30x) resulted in disrupted arborization and shortening of the motor neuron axons (Fig 5G and H). Overall, these results indicate that the partial knockdown of C9ORF72 does not induce major neuronal cell death but synergizes

Figure 5. Reduced expression of C9ORF72 synergizes Ataxin-2 toxicity.

A Left panel, representative images of organotypic cultures of E18 mouse cortical neurons co-transfected with HA-tagged *ATXN2* with control (Q22x) or intermediate (Q30x) polyQ size and transduced with lentivirus expressing either control shRNA or shRNA targeting *C9orf72* mRNA. Right panel, quantification of Ataxin-2 aggregates. Scale bars, 10 μm. Nuclei were counterstained with DAPI.

B Cell viability (tetrazolium assay) of GT1-7 neuronal cells co-transfected with HA-tagged *ATXN2* with control (Q22x) or intermediate (Q30x) polyQ size and control siRNA or siRNA targeting *C9orf72* mRNA. Error bars indicate SEM, $n = 3$.

C Tracing of the swimming trajectories of 48 h post-fertilization zebrafish larvae following light touch.

D–F Quantification of the touch-evoked swimming distance (D), average velocity (E), and maximum velocity attained (F) shows significant functional impairment of the zebrafish injected with HA-tagged *ATXN2* with intermediate length of polyQ (Q30x) and antisense morpholino oligonucleotides (AMOs) against *C9orf72* compared to control HA-tagged *ATXN2* (Q22x) or to the sole injection of AMO against *C9orf72*.

G Representative images of motor neurons visualized with anti-Sv2 immunohistochemistry show severe axonopathy in fish injected with both *ATXN2* Q30x and AMO against *C9orf72* compared to zebrafish injected with control *ATXN2* Q22x or with AMO against *C9orf72* alone.

H Quantification of the motor neuron axonal length demonstrates a significant decrease in axonal length in 48 h post-fertilization zebrafish larvae injected with both *ATXN2* Q30x and AMO against *C9orf72* compared to controls.

Data information: Error bars indicate SEM. Student's *t*-test, ***$P < 0.001$ (A, B) or one-way ANOVA, **$P < 0.01$ (D–F, H), $n = 3$.

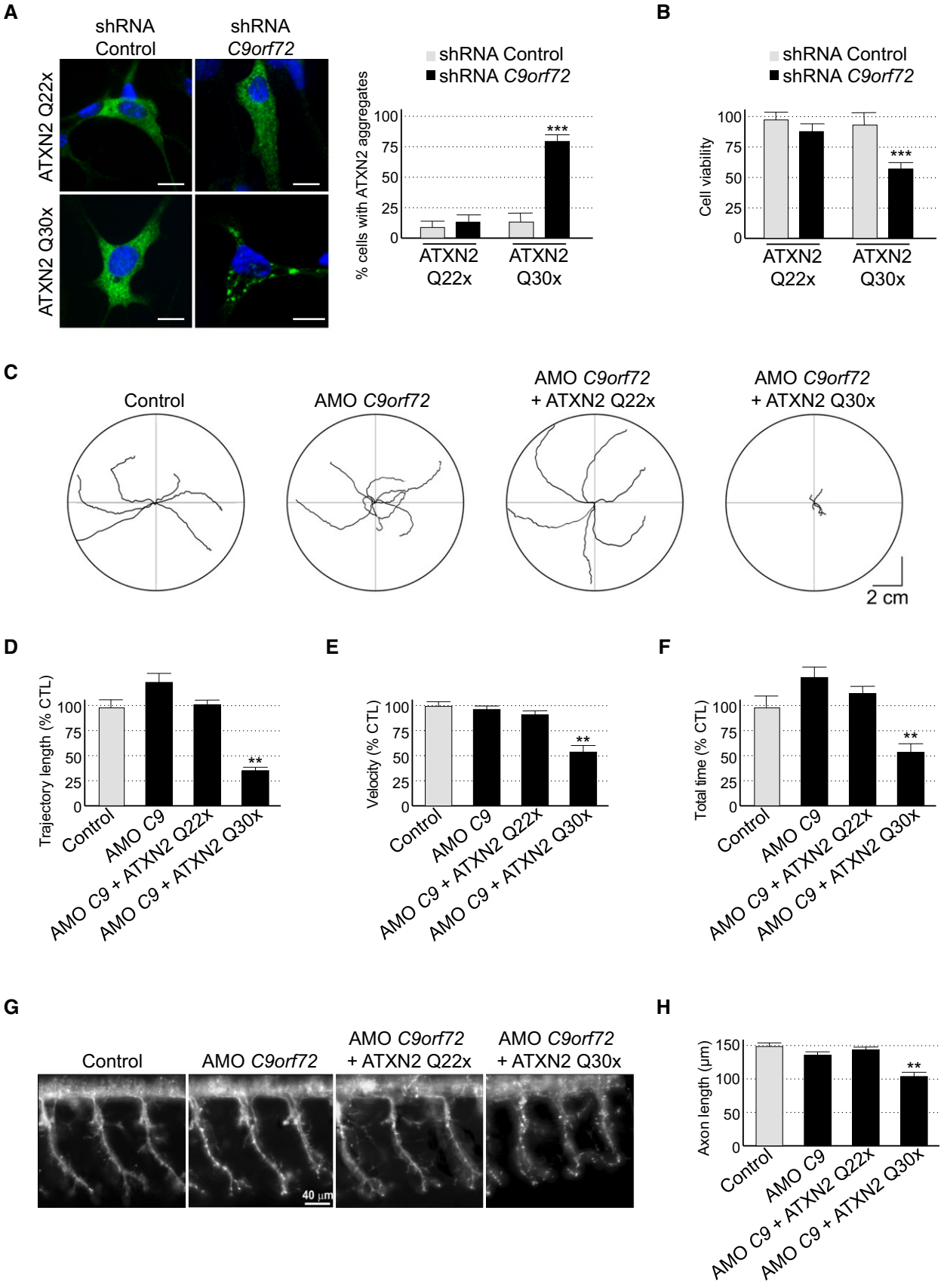


Figure 5.

the toxicity of Ataxin-2 with intermediate length of polyQ, resulting in alterations of the motor neuron and of the locomotor phenotype *in vivo*.

Discussion

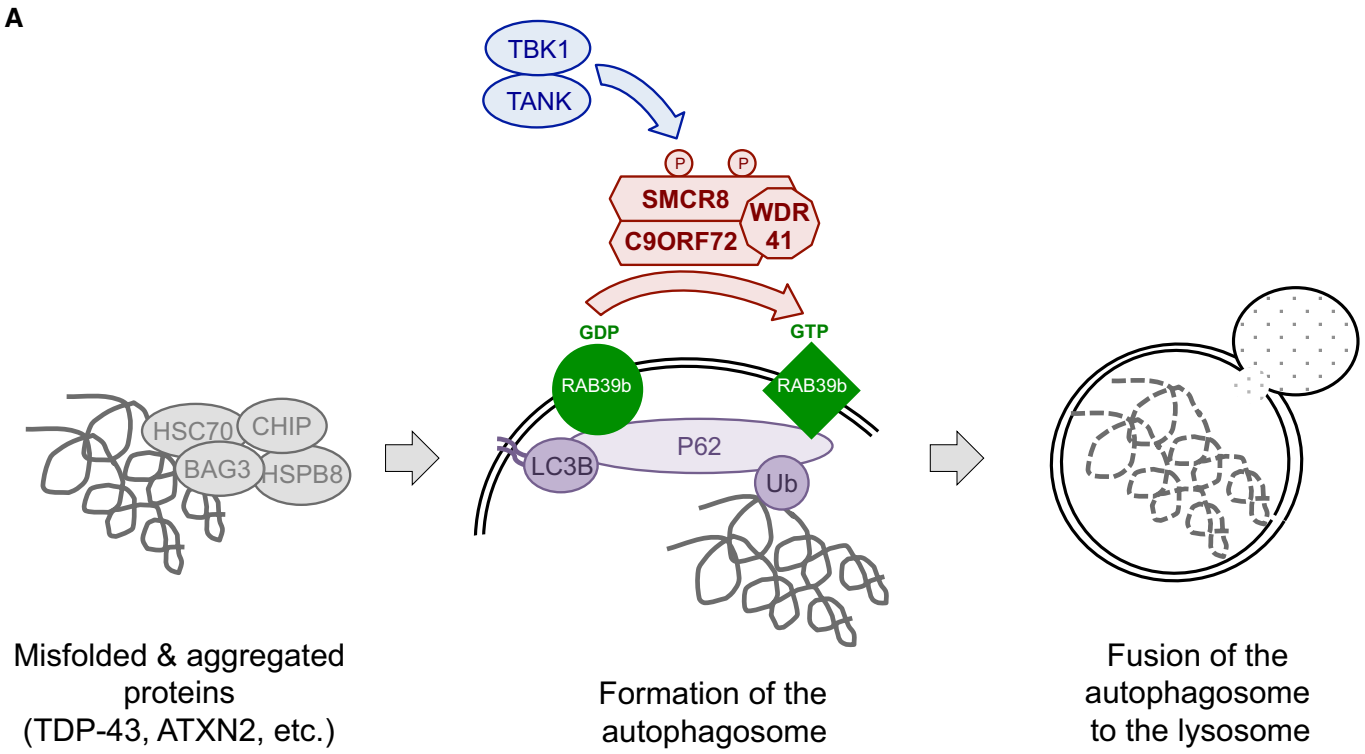
In conclusion, we found that C9ORF72 belongs to a complex containing the SMCR8 and WDR41 proteins. This complex is phosphorylated by TBK1 and ULK1 kinases, is a GDP/GTP exchange factor for RAB8a and RAB39b GTPases, and regulates autophagy (Fig 6). These results are reminiscent of the complex formed by folliculin (FLCN) and the folliculin-interacting protein-1 and folliculin-interacting protein-2 (FNIP1 and FNIP2), which contain divergent DENN modules presenting some homology with SMCR8 and C9ORF72, respectively (Zhang *et al*, 2012; Levine *et al*, 2013). FLCN is a tumor suppressor protein disrupted in various cancers and the Birt–Hogg–Dubé syndrome, which presents GTPase-activating activity for RagC/D regulating mTORC1 activity (Petit *et al*, 2013; Tsun *et al*, 2013). Also and similar to the C9ORF72 complex, FLCN is a GEF for Rab GTPases, is phosphorylated by ULK1, and is involved in the control of autophagy (Nookala *et al*, 2012; Dunlop *et al*, 2014).

However, our results indicating that C9ORF72 in complex with SMCR8 interacts with RAB8a and RAB39b are different from a previous report where C9ORF72 was found to interact with RAB1, RAB5, RAB7, and RAB11 (Farg *et al*, 2014). These discrepancies are inherent to the different approaches used in these two studies. Notably, we tested C9ORF72 in complex, while Farg and colleagues studied C9ORF72 in isolation. Similarly, composition of the immunoprecipitation washing buffers is different with 50 mM Tris in Farg and colleagues versus 50 mM Tris with 150 mM NaCl in our study. Moreover, we found that commercial antibodies against C9ORF72 were of poor specificity to carry immunoprecipitation of endogenous proteins. A second divergence between these two studies is that Farg and colleagues observed that siRNA-mediated depletion of C9orf72 increases LC3B-II levels in basal condition. In contrast, we observed little effect of siRNA- or shRNA-mediated depletion of C9ORF72 on LC3B in basal condition, but observed a significant inhibition of LC3B lipidation when autophagy flux was investigated. These differences may originate from the different degree of siRNA-mediated depletion of C9ORF72 and/or from the different cells employed (primary E18 cortical neurons and GT1-7 cells in our study versus SH-SY5Y cells in Farg and colleagues) since basal autophagy and LC3B levels are known to vary according to the cell types.

Importantly, a role of C9ORF72 in autophagy is consistent with the increased accumulation of P62 and susceptibility to inhibition of autophagy observed in human neurons derived from C9ORF72 iPS cells (Almeida *et al*, 2013). Similarly, accumulation of TDP-43- and P62-positive protein aggregates upon reduction in C9ORF72 expression is reminiscent of a dysfunction of autophagy and reproduces key histopathological features of ALS-FTD patients (Neumann *et al*, 2006; Al-Sarraj *et al*, 2011). In that aspect, it is also striking to note that various mutations causing ALS-FTD are found in genes involved in protein clearance pathways, including *UBQLN2*, *CHMP2B, VCP*, *OPTN*, *SQSTM1*, and *TBK1* (Skibinski *et al*, 2005; Johnson *et al*, 2010; Maruyama *et al*, 2010; Deng *et al*, 2011; Fecto *et al*, 2011; Cirulli *et al*, 2015; Freischmidt *et al*, 2015). Thus, our work linking decreased expression of C9ORF72 to partial deficient autophagy provides further support to compromised protein clearance mechanisms in ALS-FTD. However, it is important to note that our siRNA approach resulted in loss of 75% or more of *C9orf72* expression, while C9ORF72 levels are reduced by 50% or less in brain of individuals with ALS-FTD. Similarly, mutations in *TBK1* cause disease through a haploinsufficiency mechanism, while our siRNA reduced Tbk1 expression by nearly 80%. Thus, we can only speculate that partial reduction in C9ORF72 or TBK1 activity in patients may result in a suboptimal autophagy pathway, which in turn may contribute to disease pathogenesis.

A second significant conclusion of this study is that constitutively active GTP-locked RAB39b corrects autophagy alteration caused by either loss of TBK1 or C9ORF72. Thus, we propose that TBK1, C9ORF72 complex, and RAB39b belong to a common pathway regulating autophagy in neuronal cells (Fig 6A). This model is similar to the ULK1-mediated phosphorylation of the guanine nucleotide exchange factor DENND3 that activates Rab12 and promotes autophagy (Xu *et al*, 2015), suggesting that phosphorylation of GDP/GTP exchange factors may be a widespread and novel pathway to activate Rab GTPase in autophagy. Since RAB39b and RAB8a interact with P62 and OPTN, our results also support a model where autophagy receptors, such as P62 or OPTN, function as essential hubs to gather autophagy substrates with LC3B but also with specific Rab GTPases and their GEF effectors and kinase regulators in order to initiate autophagy precisely at the site of protein aggregates, dysfunctional organelles, or intracellular pathogens (Fig 6B). This model is supported by the interaction and phosphorylation of P62 or OPTN by TBK1 in xenophagy (Wild *et al*, 2011; Pilli *et al*, 2012), but also by the recent reports of OPTN and TBK1 importance for the PINK1-Parkin mitophagy pathway, which is altered in Parkinson's disease (Heo *et al*, 2015; Lazarou *et al*, 2015; Matsumoto *et al*, 2015). In that aspect, mutations in the X-linked *RAB39b* gene lead to early-onset Parkinson's disease (Wilson *et al*, 2014), while atypical parkinsonism has been observed in rare cases of individuals with GGGGCC expansion in *C9ORF72* (Wilke *et al*, 2016).

This work also raises several questions. First, we noted that the C9ORF72 complex acts as a GEF toward RAB8a and RAB39b; however, we did not test all existing Rab proteins and the C9ORF72 complex may potentially regulate various other Rab GTPases. Similarly, it is unclear which Rab proteins initiate and regulate autophagy in which tissue, developmental time, or conditions. In that aspect, the precise molecular functions of RAB8 and RAB39 in autophagy remain also to be elucidated. Binding of RAB8a and RAB39b to P62 suggest that these Rab GTPases may act on formation of the autophagosome. However, whether these Rab GTPases can also regulate the transport or fusion of the autophagosome to multivesicular body or lysosome is to be determined. Also, we found that phosphorylation of SMCR8 by TBK1 is important to control autophagy in neuronal cells. However, it remains to test the importance of SMCR8 phosphorylation on the GDP/GTP exchange activity of the C9ORF72 complex. Likewise, the signaling pathways activating TBK1 is yet unclear. Indeed, TBK1 activation may require an upstream kinase (Heo *et al*, 2015), while a second and non-exclusive model proposes that local concentration of TBK1 through its recruitment via OPTN may auto-activate TBK1 through trans-autophosphorylation (Matsumoto *et al*, 2015). Furthermore, SMCR8



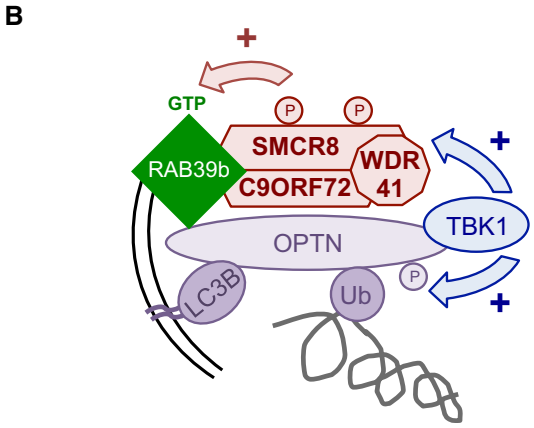


Figure 6. Tentative model of C9ORF72 function.

A C9ORF72 in complex with SMCR8 and WDR41 acts as a GDP/GTP exchange factor for RAB39b GTPase, which interacts with the P62 autophagy receptor. Phosphorylation of SMCR8 by TBK1 potentially promotes C9ORF72 GEF activity and enhances autophagy turnover of proteins such as TDP-43 or Ataxin-2 with intermediate polyQ size. In the absence of C9ORF72, autophagy clearance of these proteins is reduced and TDP-43, P62, or Ataxin-2 with intermediate length of polyQ accumulates into cytoplasmic aggregates.

B Autophagy receptors such as P62 or OPTN act as hubs to gather Rab GTPases with their GEF effectors and kinase regulators to initiate autophagy precisely at the site of damaged organelles, protein aggregates, or intracellular pathogens.

is also interacting with proteins of the ULK1, mTOR, and AMPK kinase complexes (Table EV1) and SMCR8 is reported phosphorylated by mTOR (Hsu *et al*, 2011) and AMPK (Hoffman *et al*, 2015; Schaffer *et al*, 2015), but with unknown consequences.

Also intriguing is the accumulation of TDP-43 aggregates upon depletion of C9ORF72 in primary culture of E18 cortical mouse neurons but not in GT1-7 cells. This discrepancy may originate from the reduced time frame (24–48 h) of analysis in GT1-7 cells compared to primary neuronal culture transduced with shRNA lentivirus for 7 days. In support of our work, a causal link between altered autophagy and accumulation of cytoplasmic aggregates of TDP-43 is long established (Filimonenko *et al*, 2007; Ju *et al*, 2009; Urushitani *et al*, 2010; Wang *et al*, 2012; Barmada *et al*, 2014; Scotter *et al*, 2014) and is consistent with the TDP-43-positive

neuronal inclusions observed in ALS-FTD patients with mutation in genes involved in autophagy or in protein clearance pathway, including *UBQLN2*, *TBK1*, *OPTN*, *SQSTM1*, *VCP*, or *GRN* (Hu *et al*, 2010; Brady *et al*, 2013; review in Majcher *et al*, 2015 and in Taylor, 2015). However, this questions whether the aggregates of TDP-43 observed in the vast majority of sporadic ALS-FTD cases are also caused by some deficiency in autophagy and/or protein clearance mechanisms, and if so, what would be the underlying pathogenic mechanisms. In that aspect, recent evidences of accumulation of TDP-43 aggregates in aging human brain (Uchino *et al*, 2015) might be related to the known decline of autophagy with age (Sun *et al*, 2015).

Finally, we found that partial loss of C9ORF72 promoted accumulation of P62 in aggregates, but had only mild effect on LC3B levels. Similarly, reduced expression of C9orf72 had little effect on neuronal cell viability. This is consistent with the absence of neurodegenerative phenotypes observed in mouse depleted of *C9orf72* expression (Lagier-Tourenne *et al*, 2013; Koppers *et al*, 2015). These results may suggest that either the activity of C9ORF72 is redundant with other proteins or that C9ORF72 is not crucial for basal autophagy but may play a more restricted role in a specific autophagy subpathway. In support of that latter hypothesis, we found that decreased expression of C9ORF72 promotes specifically the aggregation of Ataxin-2 with intermediate length of polyQ. These results support the co-incidence of expanded GGGGCC repeats in *C9ORF72* with intermediate length of polyQ repeats in Ataxin-2 in ALS-FTD patients (Elden *et al*, 2010; Daoud *et al*, 2011; Ross *et al*, 2011; Van Damme *et al*, 2011; Lattante *et al*, 2014). However, why the decreased expression of C9ORF72 promotes the aggregation and toxicity of Ataxin-2 with polyQ expansion, but not the aggregation of Ataxin-3 or Huntingtin with polyQ expansion, is intriguing and the cause of this specificity remains to be explored. In that aspect, both TDP-43 and Ataxin-2 are RNA-binding proteins, thus questioning whether a selective autophagy of RNA-protein granules (RNAphagy) is controlled by C9ORF72 and altered specifically in ALS-FTD (Buchan *et al*, 2013; Fujiwara *et al*, 2013). It is also possible that we missed the deleterious effect of C9ORF72 loss on other polyQ proteins due to the reduced time frame of our study or to other inherent limitations of cell cultures. Similarly, it is possible that Ataxin-2 with a control length (22Q) of polyglutamine might form microaggregates that would be toxic on a longer time period of analysis. This hypothesis is sustained by the observation that Ataxin-2 with both normal and intermediate length of polyQ synergizes the toxicity of TDP-43 in fly, however with normal polyQ size being less toxic than intermediate polyQ length (Kim *et al*, 2014). Thus, murine models would be instrumental to test the pathological consequences of loss of C9ORF72 in the presence of Ataxin-2 or other proteins with polyglutamine expansion. It also remains to test whether decreased expression of C9ORF72 may synergize toxicity of other stress, notably some that are inherent to the expanded GGGGCC repeats such as accumulation of GGGGCC RNA foci and/or RAN-translated DPRs. This hypothesis is particularly appealing in light of the association of C9ORF72 mRNA expression with patient survival (van Blitterswijk *et al*, 2015) and the recent reports of no overt neurodegenerative phenotype in BAC transgenic mouse models with normal expression of C9orf72 but overexpression of expanded GGGGCC repeats. These mice present RNA foci and aggregates of dipeptide repeat proteins but develop only subtle behavioral phenotypes (O'Rourke *et al*, 2015; Peters *et al*, 2015). Similarly, mice depleted of *C9orf72* expression in brain or in neurons present no overt neurodegenerative phenotypes (Lagier-Tourenne *et al*, 2013; Koppers *et al*, 2015). These results suggest that the sole loss of C9ORF72 or the expression of GGGGCC RNA and DPR in isolation is not sufficient to be pathogenic, but it remains to test whether a reduced expression of C9ORF72 would synergize the toxicity of GGGGCC RNA or DPRs. This synergic model is consistent with the absence of ALS/FTD patients with null alleles or missense mutations in *C9ORF72*, as well as by increasing genetic evidences of oligogenicity in ALS-FTD (Ferrari *et al*, 2012; Van Blitterswijk *et al*, 2012; van Blitterswijk *et al*, 2013; Cady *et al*, 2015; Lattante *et al*, 2015a; Pottier *et al*, 2015).

In conclusion, our results support a double-hit mechanism in ALS-FTD, where the sole decreased expression of C9ORF72 does not explain alone the pathogenicity of the expanded GGGGCC repeats but may synergize other stress such as Ataxin-2 with intermediate polyQ length, while association with other stress such as GGGGCC RNA foci or Ran-translated DPRs remains to be formally tested. Finally, if loss of C9ORF72 leads to partial dysfunction of autophagy, one may hope that pharmacological compounds activating autophagy may contribute to alleviate some pathological features of ALS-FTD.

Materials and Methods

Constructions

PCMV6 containing C-terminally Flag-tagged human cDNAs of SQSTM1 (P62), OPTN, SMCR8, WDR41, ULK1, NAP1, SINTBAD, TANK, and RAB GTPases were purchased from OriGene. Optimized cDNAs for human N-terminally HA-tagged *C9ORF72* and Flag-tagged *TBK1* cloned into pcDNA3 were purchased from GenScript. Human cDNAs of *ATXN2* with 22 or 30 glutamines were fused to an N-terminal HA tag and cloned into pcDNA3. Constitutively active Q68L and negative S22N mutants of N-terminally HA-tagged RAB39b were constructed by inverse PCR. Similarly, mutations of C-terminally HA-tagged SMCR8 phosphorylation sites (S400, S402, S492, S562, T666, and T796D) in alanine or aspartate were constructed by inverse PCR.

HA-Flag tandem affinity purification

12×10^6 Neuro-2A cells were transfected with 48 μg of either control- or C9ORF72-Flag-Ha plasmid using Fugene HD (Promega) for 24 h, and proteins were purified by Ha-Flag tandem purification kit according to the manufacturer's instruction (Sigma-Aldrich). The bound proteins were visualized by silver staining (SilverQuest, Invitrogen) after separation on a 4–12% bis-Tris Gel (NuPAGE), and interactant proteins were identified using NanoESI_Ion Trap (Thermo Fisher).

Immunofluorescence

Coverslips were incubated for 10 min in PBS with 4% paraformaldehyde, washed with PBS, and incubated in PBS plus 0.5% Triton X-100 for 10 min. The cells were washed three times with PBS, and

the coverslips were incubated for 1 h with primary antibody against the HA tag (26183, Pierce). P62/Sqstm1 (ab56416, Abcam) or anti-phospho-TDP-43 (pS409/410, Cosmo Bio). After washing with PBS, the coverslips were incubated with goat anti-mouse secondary antibody conjugated with Alexa 488 (Interchim SA) for 1 h, washed twice with PBS, and incubated for 2 min in PBS/DAPI (1/10,000 dilution). Coverslips were rinsed twice before mounting in Pro-Long media (Molecular Probes) and were examined using confocal microscope.

Immunoprecipitation

6.25×10^5 HEK293 cells (1 well of a 6-well plate) were co-transfected for 24 h with 1 μg of plasmids expressing HA-tagged cDNA and 1 μg of plasmids expressing Flag-tagged cDNA using Fugene HD (Promega). Cells were scraped into RIPA buffer (50 mM Tris–HCI pH 7.6, 150 mM NaCl, 1% NP-40) and centrifuged for 15 min at 18,000 *g* at 4°C; 20 μl of pre-washed HA magnetic beads (Dynabeads) was added, and immunoprecipitation was carried for 1 h at 4°C with constant rotation. After three washes with 50 mM Tris–HCI pH 7.6, 150 mM NaCl, 0.05% Tween, bound proteins were eluted in SDS–PAGE loading buffer and analyzed by Western blot using antibodies against the Flag tag (PA1-984B, Pierce) or the HA tag (26183, Pierce). For endogenous immunoprecipitation, SMCR8 1D2 mouse monoclonal antibody was incubated with mouse brain extract overnight in RIPA buffer. Pre-cleared A/G magnetic beads (Life Technologies) were added, and immunoprecipitation was carried out for 1 h at 4°C with constant rotation. The beads were washed three times with 50 mM Tris–HCI, 150 mM NaCl, 0.05% Tween and then boiled at 95°C for 5 min in SDS–PAGE loading buffer. Bound proteins were analyzed by Western blot using antibodies against SMCR8 (1D2), C9ORF72 (22637-1-AP, Proteintech), WDR41 (NBP1-83812, Novus Biological), Rab8a (11792-1-AP, Proteintech), Rab39b (12162-1-AP, Proteintech), Rab5 (11671-1-AP, Proteintech), Rab7 (Ab137029, Abcam).

In vitro GDP/GTP exchange assay

About 40 pmol (~2 μg) of purified recombinant GST-RAB protein was loaded with 1 μCurie of α^{32}P-labeled GDP (Hartmann Analytic) in 10 μl of buffer assay (150 mM NaCl, 50 mM Hepes, 1 mg/ml BSA, 2.5 mM EDTA, pH 7.5) for 30 min at 30°C. Increased quantities of recombinant purified C9ORF72 or C9ORF72/SMCR8/WDR41 complex were added and incubated for further 30 min at 30°C. Twenty microliters of pre-washed GST magnetic beads (Dynabeads) was added, and pull-down was carried out for 30 min at 4°C. After three washes in reaction buffer, radioactivity was measured by using a scintillation counter (Beckman Coulter).

In vitro phosphorylation

Eight micrograms of recombinant purified C9ORF72/SMCR8/WDR41 complex was incubated for 30 min at 30°C with 10 μCurie of γ^{32}P-labeled ATP in 20 μl of kinase buffer assay (150 mM NaCl, 20 mM HEPES, 2 mM $MgCl_2$, 25 mM β-glycerophosphate, 100 μM orthovanadate, pH 7.5) with or without 2 μg of recombinant purified ULK1 or TBK1 protein (OriGene). The reaction was stopped by the addition of SDS–PAGE loading buffer, boiled for 3 min at 90°C, and run on 4–12% bis-Tris Gel (NuPAGE). The gel was then either used for Western blotting assay or dried and imaged (Typhoon scanner, GE Healthcare).

Neuronal cell cultures, transfection, and treatments

Primary cortical neurons were prepared from C57Bl/6 E18 embryos and grown on poly-L-lysine-coated 24-well plates in neurobasal medium (NBM) supplemented with 1× B27, 0.5 mM L-glutamine, and 100 IU/ml penicillin/streptomycin at 37°C with 5% CO_2. Neurons were transduced at Day 3 with recombinant lentivirus expressing either control shRNA or shRNA against *C9orf72* (SK02-040236-00-10 SMARTvector 2.0 hEF1a Lentiviral Mouse 3110043O21Rik shRNA, Thermos). After overnight incubation, lentivirus was removed and fresh media were added. After 5–7 days, neurons were analyzed by immunofluorescence or by Western blot analysis. GT1-7 cells were grown in 10% fetal bovine serum, gentamicin, and penicillin at 37°C in 5% CO_2, plated in DMEM and 0.1% fetal bovine serum, and transfected for 24–48 h using Lipofectamine 2000 (Fisher Scientific) and/or RNAimax (Fisher Scientific) with siRNA control or targeting *C9orf72*, *Smcr8* 3′UTR, or *Tbk1* (ON-TARGETplus, Dharmacon). Neurons were treated with either 10 μM of rapamycin (Millipore) for 15 h or 100 nM of bafilomycin (Sigma) for 15 h or 250 nM of Torin-1 (Tocris) for 2 h before analysis.

Recombinant protein production and purification

For RAB GTPases, *E. coli* BL21(RIL) pRARE competent cells (Invitrogen) were transformed with pet28-GST-RAB GTPase vectors, grown at 37°C in 400 ml of LB medium supplemented with kanamycin until OD_{600} = 0.5, 0.5 mM IPTG was added and the culture was further incubated for 4 h at 30°C. Harvested cells were sonicated in 300 mM NaCl, 50 mM Tris–Cl pH 7.5, 1 mM DTT, 5 mM EDTA and centrifuged for 20 min at 20,000 *g* and recombinant GST-RAB GTPase proteins were purified using the GST purification Kit (Novagen), quantified, dialyzed, and stored in 150 mM NaCl, 20 mM HEPES, 2 mM $MgCl_2$, 20% glycerol. Concerning C9ORF72 complex, 2×10^6 of SF9 (*Spodoptera frugiperda*) cells (one T25 flask) were co-transfected with 500 ng of *Bsu36I*-linearized BAC10:KO1629 DNA and 2 μg of pMF-Dual vectors containing either HIS-C9ORF72, SMCR8, or WDR41 and incubated at 27°C for 6 days. Harvested baculovirus were then tested, amplified and used to infect 2 l of SF9 cell culture for protein production and purification using the HIS purification kit (Novagen) with sonication and washing in 500 mM NaCl, 50 mM Tris–Cl pH 7.5, 50 mM imidazole, elution with 150 mM NaCl, 50 mM Tris–Cl pH 7.5, 1 mM DTT, 5 mM EDTA, 200 mM imidazole, dialysis, and storage in 150 mM NaCl, 20 mM HEPES, 2 mM $MgCl_2$, 20% glycerol.

Monoclonal antibody production

To generate anti-SMCR8 monoclonal antibodies, 8-week-old female BALB/c mice were injected intraperitoneally with 100 μg of recombinant purified SMCR8:C9ORF72 complex and 200 μg of poly (I/C) as adjuvant. Three injections were performed at 2-week intervals, and 4 days prior to hybridoma fusion, mice with positively reacting sera were re-injected. Spleen cells were fused with

Sp2/0.Agl4 myeloma cells as described by De StGroth and Scheidegger (1980). Hybridoma culture supernatants were tested at Day 10 by ELISA for cross-reaction with recombinant purified C9ORF72: SMCR8 complex. Positive supernatants were then tested by immunofluorescence and Western blot on HA-tagged SMCR8 and HA-tagged C9ORF72 transfected COS-1 cells. Specific cultures were cloned twice on soft agar. Specific hybridomas were established, and ascites fluid was prepared by injection of 2×10^6 hybridoma cells into Freund adjuvant-primed BALB/c mice. All animal experimental procedures were performed according to the European authority guidelines.

Western blotting

Proteins were denatured for 3 min at 95°C, separated on 4–12% bis-Tris Gel (NuPAGE), transferred on nitrocellulose membranes (Whatman Protran), blocked with 5% non-fat dry milk in Tris-buffered saline (TBS) buffer, incubated with anti-Flag (PA1-984B, Pierce), HA (26183, Pierce), C9ORF72 (22637-1-AP, Proteintech), LC3B (ab51520, Abcam), GAPDH (ab125247, Abcam), TDP-43 (3449S, Cell Signaling), SMCR8 (1D2, 1/200) in TBS plus 5% non-fat dry milk, washed three times, and incubated with anti-rabbit or anti-mouse peroxidase antibody (1:3,000, Cell Signaling) for 1 h in TBS, followed by washing and ECL chemoluminescence revelation (Amersham ECL Prime).

Zebrafish studies

Adult and larval zebrafish (*Danio rerio*) were maintained at the ICM fish facility and bred according to the National and European Guidelines for Animal Welfare. Experiments were performed on wild-type embryos from AB and TL strains. Hb9:GFP Tg(Mnx1:GFP) transgenic lines were used to label motor neurons and their axonal projections. All procedures for zebrafish experimentation were approved by the Institutional Ethics Committee at the Research Center of the ICM and by French and European legislation. The *ATXN2* constructs were used in all the experiments described here at the final DNA concentration of 75 ng/µl. Embryos were maintained at 28°C and manually dechorionated using fine forceps at 24 hpf. Only zebrafish without developmental abnormalities (qualified as malformed) and the swimming trajectories at 48 h post-fertilization were selected and their swimming trajectories appropriately traced. For the percentage analysis, behavioral analysis, and axonal projections, more than three independent experiments were performed for each of the conditions described here. Antisense morpholino oligonucleotides (AMOs) were designed complementary to bind to an upstream ATG that would block both transcripts of the zebrafish *C9orf72* (C9orf72) and synthesized from GeneTools. The *C9orf72* AMO sequence was ATTGTGGAGGACAGGCTGAAGACAT and known to bind to the following sequence in the *C9orf72* mRNA: [(ATG)TCTTCAGCCTGTCCTCCACAAT]. A control AMO (mismatch), containing five mismatch nucleotides with the C9orf72 AMO sequence and not binding anywhere in the zebrafish genome, was used to assess the specificity of the observed phenotype (ATTcTcGAGcACAGcCTcAAGACAT). For the genetic interaction experiments, microinjections were performed at 0.2 mM for C9orf72-AMO and 0.2 mM for C9orf72-mis. At these concentrations, C9orf72 and mismatch AMOs do not lead to phenotypic features associated with deficient swimming or morphological abnormalities (Ciura *et al*, 2013; Lattante *et al*, 2015a,b). For the percentage analysis, behavioral analysis, and axonal projections, more than three independent experiments were performed for each of the conditions including co-injections of *ATXN2* Q22x and *ATXN2* Q30x alongside C9orf72 and mismatch AMOs. Zebrafish embryos at 48 hpf were analyzed for any morphological abnormalities and touched lightly at the level of the tail with a pipette tip with their locomotor behavior scored. Thus, for each injection set, larvae and embryos were separated into the following groups: dead, curly, and monster groups (developmentally aberrant fish). TEER episodes were performed only in zebrafish that appeared morphologically normal (normal TEER observed) and were recorded for each of the conditions with a Grasshopper 2 Camera (Point Grey Research) at 30 Hz. The videos were then analyzed using the manual tracking plugin of ImageJ 1.45r software, and the swim duration, swim distance, and maximum swim velocity of the fish were calculated as previously described (Ciura *et al*, 2013; Lattante *et al*, 2015b). To correlate gene expression with cell morphology, the axonal projections of motor neurons in selected HB9 zebrafish embryos GFP-positive at 48 hpf. In addition, axonal projections were labeled using the synaptic vesicle marker, SV2 as previously described (Kabashi *et al*, 2010). Fluorescent images of fixed embryos were taken using the fluorescence Automated Inverted Microscope System Olympus IX83 equipped with a Hamamatsu ORCA-flash 2.8 digital camera. Image acquisition was performed with the Olympus cellSens software. Axonal projections from primary motor neurons at a defined location in the intersomitic segments were determined. Analysis of Z-stacks by fluorescence microscopy was performed in three to four axonal projections per animal. The axonal length to the first branching was determined by tracing the labeled axon from the spinal cord to the point where it branches using ImageJ. These values were averaged for each of the animals analyzed (minimum 15 zebrafish per condition) for the various conditions. Primers for RT–qPCR quantification of HA-tagged *ATXN2* injected in zebrafish are F-TGGTTCTCCAGCTCCTGTCT and R-TGACCACTGATGACCACGTT. *ATXN2* levels were normalized to endogenous *Gapdh* (QuantiTect Primer Assay, Dr_gapdh_1_SG).

Statistical analysis

All cell experiments are represented as average ± standard error of mean (SEM) with significance determined using Student's *t*-test. All data values for the zebrafish experiments are represented as average ± standard error of mean (SEM) with significance determined using one-way ANOVAs.

Expanded View for this article is available online.

Acknowledgements

We thank Pamela Mellon (UCSD, USA) for the gift of the GT1-7 cells; Terje Johansen (University of Tromsø, Norway) for the gift of the P62, OPTN, and LC3B plasmids; Jochen Weishaupt (Ulm University, Germany) for the gift of the TBK1 vector; and Aaron Gitler (Stanford University School of Medicine, USA) and Daisuke Ito (Keio University, Tokyo, Japan) for the gift of the Ataxin-2 plasmids. This work was supported by Fondation de France Thierry Latran #57486 "Model-ALS", AFM grant #18605 "Role of C9ORF72 in ALS-FTD", ERC-2012-StG #310659 "RNA DISEASES", ANR-10-LABX-0030-INRT, and

ANR-10-IDEX-0002-02 (NCB); Atip/Avenir from Inserm, Career Integration Grant (Marie Curie Actions), Robert Packard Foundation, E-rare ERA-NET program, AFM, ARSLA, France-Alzheimer Association, and the program "Investissements d'avenir" ANR-10-IAIHU-06 (EK). SC is supported by an AFM postdoctoral fellowship and M-LC by a postdoctoral fellowship from the Fondation Cognacq-Jay.

Author contributions

Experiments were performed by CS, MLC, AG, IKC, CJC, MOA, FR and SC. Data were collected and analyzed by CS, AP, SC, EK, and NCB. The study was designed, coordinated, and written by CS, SC, EK, and NCB.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Almeida S, Gascon E, Tran H, Chou HJ, Gendron TF, Degroot S, Tapper AR, Sellier C, Charlet-Berguerand N, Karydas A, Seeley WW, Boxer AL, Petrucelli L, Miller BL, Gao FB (2013) Modeling key pathological features of frontotemporal dementia with C9ORF72 repeat expansion in iPSC-derived human neurons. *Acta Neuropathol* 126: 385 – 399

Al-Sarraj S, King A, Troakes C, Smith B, Maekawa S, Bodi I, Rogelj B, Al-Chalabi A, Hortobágyi T, Shaw CE (2011) P62 positive, TDP-43 negative, neuronal cytoplasmic and intranuclear inclusions in the cerebellum and hippocampus define the pathology of C9orf72-linked FTLD and MND/ALS. *Acta Neuropathol* 122: 691 – 702

Arai T, Hasegawa M, Akiyama H, Ikeda K, Nonaka T, Mori H, Mann D, Tsuchiya K, Yoshida M, Hashizume Y, Oda T (2006) TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Biochem Biophys Res Commun* 351: 602 – 611

Arndt V, Dick N, Tawo R, Dreiseidler M, Wenzel D, Hesse M, Fürst DO, Saftig P, Saint R, Fleischmann BK, Hoch M, Höhfeld J (2010) Chaperone-assisted selective autophagy is essential for muscle maintenance. *Curr Biol* 20: 143 – 148

Ash PE, Bieniek KF, Gendron TF, Caulfield T, Lin WL, van Blitterswijk MM, Jansen-West K, Paul JW 3rd, Rademakers R, Boylan KB, Dickson DW, Petrucelli L (2013) Unconventional translation of C9ORF72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. *Neuron* 77: 639 – 646

Barmada SJ, Serio A, Arjun A, Bilican B, Daub A, Ando DM, Tsvetkov A, Pleiss M, Li X, Peisach D, Shaw C, Chandran S, Finkbeiner S (2014) Autophagy induction enhances TDP43 turnover and survival in neuronal ALS models. *Nat Chem Biol* 10: 677 – 685

Behrends C, Sowa ME, Gygi SP, Harper JW (2010) Network organization of the human autophagy system. *Nature* 466: 68 – 76

van Blitterswijk M, Baker MC, DeJesus-Hernandez M, Ghidoni R, Benussi L, Finger E, Hsiung GY, Kelley BJ, Murray ME, Rutherford NJ, Brown PE, Ravenscroft T, Mullen B, Ash PE, Bieniek KF, Hatanpaa KJ, Karydas A, Wood EM, Coppola G, Bigio EH *et al* (2013) C9ORF72 repeat expansions in cases with previously identified pathogenic mutations. *Neurology* 81: 1332 – 1341

van Blitterswijk M, Gendron TF, Baker MC, DeJesus-Hernandez M, Finch NA, Brown PH, Daughrity LM, Murray ME, Heckman MG, Jiang J, Lagier-Tourenne C, Edbauer D, Cleveland DW, Josephs KA, Parisi JE, Knopman DS, Petersen RC, Petrucelli L, Boeve BF, Graff-Radford NR *et al* (2015) Novel clinical associations with specific C9ORF72 transcripts in patients with repeat expansions in C9ORF72. *Acta Neuropathol* 130: 863 – 876

Bose JK, Huang CC, Shen CK (2011) Regulation of autophagy by neuropathological protein TDP-43. *J Biol Chem* 286: 44441 – 44448

Brady OA, Zheng Y, Murphy K, Huang M, Hu F (2013) The frontotemporal lobar degeneration risk factor, TMEM106B, regulates lysosomal morphology and function. *Hum Mol Genet* 22: 685 – 695

Buchan JR, Kolaitis RM, Taylor JP, Parker R (2013) Eukaryotic stress granules are cleared by autophagy and Cdc48/VCP function. *Cell* 153: 1461 – 1474

Cady J, Allred P, Bali T, Pestronk A, Goate A, Miller TM, Mitra RD, Ravits J, Harms MB, Baloh RH (2015) Amyotrophic lateral sclerosis onset is influenced by the burden of rare variants in known amyotrophic lateral sclerosis genes. *Ann Neurol* 77: 100 – 113

Chan EY, Kir S, Tooze SA (2007) siRNA screening of the kinome identifies ULK1 as a multidomain modulator of autophagy. *J Biol Chem* 282: 25464 – 25474

Chew J, Gendron TF, Prudencio M, Sasaguri H, Zhang YJ, Castanedes-Casey M, Lee CW, Jansen-West K, Kurti A, Murray ME, Bieniek KF, Bauer PO, Whitelaw EC, Rousseau L, Stankowski JN, Stetler C, Daughrity LM, Perkerson EA, Desaro P, Johnston A *et al* (2015) Neurodegeneration. C9ORF72 repeat expansions in mice cause TDP-43 pathology, neuronal loss, and behavioral deficits. *Science* 348: 1151 – 1154

Cirulli ET, Lasseigne BN, Petrovski S, Sapp PC, Dion PA, Leblond CS, Couthouis J, Lu YF, Wang Q, Krueger BJ, Ren Z, Keebler J, Han Y, Levy SE, Boone BE, Wimbish JR, Waite LL, Jones AL, Carulli JP, Day-Williams AG *et al* (2015) Exome sequencing in amyotrophic lateral sclerosis identifies risk genes and pathways. *Science* 347: 1436 – 1441

Ciura S, Lattante S, Le Ber I, Latouche M, Tostivint H, Brice A, Kabashi E (2013) Loss of function of C9orf72 causes motor deficits in a zebrafish model of Amyotrophic Lateral Sclerosis. *Ann Neurol* 74: 180 – 187

Cooper-Knock J, Walsh MJ, Higginbottom A, Robin Highley J, Dickman MJ, Edbauer D, Ince PG, Wharton SB, Wilson SA, Kirby J, Hautbergue GM, Shaw PJ (2014) Sequestration of multiple RNA recognition motif-containing proteins by C9orf72 repeat expansions. *Brain* 137 : 2040 – 2051

Daoud H, Belzil V, Martins S, Sabbagh M, Provencher P, Lacomblez L, Meininger V, Camu W, Dupré N, Dion PA, Rouleau GA (2011) Association of long ATXN2 CAG repeat sizes with increased risk of amyotrophic lateral sclerosis. *Arch Neurol* 68: 739 – 742

De StGroth SF, Scheidegger D (1980) Production of monoclonal antibodies: strategy and tactics. *J Immunol Methods* 35: 1 – 21

DeJesus-Hernandez M, Mackenzie IR, Boeve BF, Boxer AL, Baker M, Rutherford NJ, Nicholson AM, Finch NA, Flynn H, Adamson J, Kouri N, Wojtas A, Sengdy P, Hsiung GY, Miller BL, Dickson DW, Boylan KB, Graff-Radford NR, Rademakers R (2011) Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron* 72: 245 – 256

Deng HX, Chen W, Hong ST, Boycott KM, Gorrie GH, Siddique N, Yang Y, Fecto F, Shi Y, Zhai H, Jiang H, Hirano M, Rampersaud E, Jansen GH, Donkervoort S, Bigio EH, Brooks BR, Ajroud K, Sufit RL, Haines JL *et al* (2011) Mutations in UBQLN2 cause dominant X-linked juvenile and adult-onset ALS and ALS/dementia. *Nature* 477: 211 – 215

Donnelly CJ, Zhang PW, Pham JT, Heusler AR, Mistry NA, Vidensky S, Daley EL, Poth EM, Hoover B, Fines DM, Maragakis N, Tienari PJ, Petrucelli L, Traynor BJ, Blackshaw S, Sattler R, Rothstein JD (2013) RNA toxicity from the ALS/FTD C9ORF72 expansion is mitigated by antisense intervention. *Neuron* 80: 415 – 428

Dunlop EA, Seifan S, Claessens T, Behrends C, Kamps MA, Rozycka E, Kemp AJ, Nookala RK, Blenis J, Coull BJ, Murray JT, van Steensel MA, Wilkinson S,

Tee AR (2014) FLCN, a novel autophagy component, interacts with GABARAP and is regulated by ULK1 phosphorylation. *Autophagy* 10: 1749 – 1760

Elden AC, Kim HJ, Hart MP, Chen-Plotkin AS, Johnson BS, Fang X, Armakola M, Geser F, Greene R, Lu MM, Padmanabhan A, Clay-Falcone D, McCluskey L, Elman L, Juhr D, Gruber PJ, Rüb U, Auburger G, Trojanowski JQ, Lee VM *et al* (2010) Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS. *Nature* 466: 1069 – 1075

Farg MA, Sundaramoorthy V, Sultana JM, Yang S, Atkinson RA, Levina V, Halloran MA, Gleeson PA, Blair IP, Soo KY, King AE, Atkin JD (2014) C9ORF72, implicated in amytrophic lateral sclerosis and frontotemporal dementia, regulates endosomal trafficking. *Hum Mol Genet* 23: 3579 – 3595

Fecto F, Yan J, Vemula SP, Liu E, Yang Y, Chen W, Zheng JG, Shi Y, Siddique N, Arrat H, Donkervoort S, Ajroud-Driss S, Sufit RL, Heller SL, Deng HX, Siddique T (2011) SQSTM1 mutations in familial and sporadic amyotrophic lateral sclerosis. *Arch Neurol* 68: 1440 – 1446

Ferrari R, Mok K, Moreno JH, Cosentino S, Goldman J, Pietrini P, Mayeux R, Tierney MC, Kapogiannis D, Jicha GA, Murrell JR, Ghetti B, Wassermann EM, Grafman J, Hardy J, Huey ED, Momeni P (2012) Screening for C9ORF72 repeat expansion in FTLD. *Neurobiol Aging* 33: 1850.e1 – 1850.e11

Filimonenko M, Stuffers S, Raiborg C, Yamamoto A, Malerød L, Fisher EM, Isaacs A, Brech A, Stenmark H, Simonsen A (2007) Functional multivesicular bodies are required for autophagic clearance of protein aggregates associated with neurodegenerative disease. *J Cell Biol* 179: 485 – 500

Freibaum BD, Lu Y, Lopez-Gonzalez R, Kim NC, Almeida S, Lee KH, Badders N, Valentine M, Miller BL, Wong PC, Petrucelli L, Kim HJ, Gao FB, Taylor JP (2015) GGGGCC repeat expansion in C9orf72 compromises nucleocytoplasmic transport. *Nature* 525: 129 – 133

Freischmidt A, Wieland T, Richter B, Ruf W, Schaeffer V, Müller K, Marroquin N, Nordin F, Hübers A, Weydt P, Pinto S, Press R, Millecamps S, Molko N, Bernard E, Desnuelle C, Soriani MH, Dorst J, Graf E, Nordström U *et al* (2015) Haploinsufficiency of TBK1 causes familial ALS and fronto-temporal dementia. *Nat Neurosci* 18: 631 – 636

Fujiwara Y, Furuta A, Kikuchi H, Aizawa S, Hatanaka Y, Konya C, Uchida K, Yoshimura A, Tamai Y, Wada K, Kabuta T (2013) Discovery of a novel type of autophagy targeting RNA. *Autophagy* 9: 403 – 409

Gamerdinger M, Hajieva P, Kaya AM, Wolfrum U, Hartl FU, Behl C (2009) Protein quality control during aging involves recruitment of the macroautophagy pathway by BAG3. *EMBO J* 28: 889 – 901

Gendron TF, Bieniek KF, Zhang YJ, Jansen-West K, Ash PE, Caulfield T, Daughrity L, Dunmore JH, Castanedes-Casey M, Chew J, Cosio DM, van Blitterswijk M, Lee WC, Rademakers R, Boylan KB, Dickson DW, Petrucelli L (2013) Antisense transcripts of the expanded C9ORF72 hexanucleotide repeat form nuclear RNA foci and undergo repeat-associated non-ATG translation in c9FTD/ALS. *Acta Neuropathol* 126: 829 – 844

Giannandrea M, Bianchi V, Mignogna ML, Sirri A, Carrabino S, D'Elia E, Vecellio M, Russo S, Cogliati F, Larizza L, Ropers HH, Tzschach A, Kalscheuer V, Oehl-Jaschkowitz B, Skinner C, Schwartz CE, Gecz J, Van Esch H, Raynaud M, Chelly J *et al* (2010) Mutations in the small GTPase gene RAB39B are responsible for X-linked mental retardation associated with autism, epilepsy, and macrocephaly. *Am J Hum Genet* 86: 185 – 195

Gijselinck I, Van Langenhove T, van der Zee J, Sleegers K, Philtjens S, Kleinberger G, Janssens J, Bettens K, Van Cauwenberghe C, Pereson S, Engelborghs S, Van Dongen J, Vermeulen S, Van den Broeck M, Vaerenberg C, Mattheijssens M, Peeters K, Robberecht W, Cras P, Martin JJ et al (2012) A C9orf72 promoter repeat expansion in a Flanders-Belgian cohort with disorders of the frontotemporal lobar degeneration-amyotrophic lateral sclerosis spectrum: a gene identification study. *Lancet Neurol* 11: 54 – 65

Gitcho MA, Baloh RH, Chakraverty S, Mayo K, Norton JB, Levitch D, Hatanpaa KJ, White CL 3rd, Bigio EH, Caselli R, Baker M, Al-Lozi MT, Morris JC, Pestronk A, Rademakers R, Goate AM, Cairns NJ (2008) TDP-43 A315T mutation in familial motor neuron disease. *Ann Neurol* 63: 535 – 538

Goncalves A, Bürckstümmer T, Dixit E, Scheicher R, Górna MW, Karayel E, Sugar C, Stukalov A, Berg T, Kralovics R, Planyavsky M, Bennett KL, Colinge J, Superti-Furga G (2011) Functional dissection of the TBK1 molecular network. *PLoS ONE* 6: e23971

Haeusler AR, Donnelly CJ, Periz G, Simko EA, Shaw PG, Kim MS, Maragakis NJ, Troncoso JC, Pandey A, Sattler R, Rothstein JD, Wang J (2014) C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature* 507: 195 – 200

Hara T, Nakamura K, Matsui M, Yamamoto A, Nakahara Y, Suzuki-Migishima R, Yokoyama M, Mishima K, Saito I, Okano H, Mizushima N (2006) Suppression of basal autophagy in neural cells causes neurodegenerative disease in mice. *Nature* 441: 885 – 889

Hara T, Takamura A, Kishi C, Iemura S, Natsume T, Guan JL, Mizushima N (2008) FIP200, a ULK-interacting protein, is required for autophagosome formation in mammalian cells. *J Cell Biol* 181: 497 – 510

Hattula K, Peränen J (2000) FIP-2, a coiled-coil protein, links Huntingtin to Rab8 and modulates cellular morphogenesis. *Curr Biol* 10: 1603 – 1606.

Heo JM, Ordureau A, Paulo JA, Rinehart J, Harper JW (2015) The PINK1-PARKIN Mitochondrial Ubiquitylation Pathway Drives a Program of OPTN/NDP52 Recruitment and TBK1 Activation to Promote Mitophagy. *Mol Cell* 60: 7 – 20

Hoffman NJ, Parker BL, Chaudhuri R, Fisher-Wellman KH, Kleinert M, Humphrey SJ, Yang P, Holliday M, Trefely S, Fazakerley DJ, Stöckli J, Burchfield JG, Jensen TE, Jothi R, Kiens B, Wojtaszewski JF, Richter EA, James DE (2015) Global Phosphoproteomic Analysis of Human Skeletal Muscle Reveals a Network of Exercise-Regulated Kinases and AMPK Substrates. *Cell Metab* 22: 922 – 935

Hsu PP, Kang SA, Rameseder J, Zhang Y, Ottina KA, Lim D, Peterson TR, Choi Y, Gray NS, Yaffe MB, Marto JA, Sabatini DM (2011) The mTOR-regulated phosphoproteome reveals a mechanism of mTORC1-mediated inhibition of growth factor signaling. *Science* 332: 1317 – 1322

Hu F, Padukkavidana T, Vægter CB, Brady OA, Zheng Y, Mackenzie IR, Feldman HH, Nykjaer A, Strittmatter SM (2010) Sortilin-mediated endocytosis determines levels of the frontotemporal dementia protein, progranulin. *Neuron* 68: 654 – 667

Imbert G, Saudou F, Yvert G, Devys D, Trottier Y, Garnier JM, Weber C, Mandel JL, Cancel G, Abbas N, Dürr A, Didierjean O, Stevanin G, Agid Y, Brice A (1996) Cloning of the gene for spinocerebellar ataxia 2 reveals a locus with high sensitivity to expanded CAG/glutamine repeats. *Nat Genet* 14: 285 – 291

Johnson JO, Mandrioli J, Benatar M, Abramzon Y, Van Deerlin VM, Trojanowski JQ, Gibbs JR, Brunetti M, Gronka S, Wuu J, Ding J, McCluskey L, Martinez-Lage M, Falcone D, Battistini S, Salvi F, Spataro R, Sola P, Borghero G, Galassi G et al (2010) Exome sequencing reveals VCP mutations as a cause of familial ALS. *Neuron* 68: 857 – 864

Jovičić A, Mertens J, Boeynaems S, Bogaert E, Chai N, Yamada SB, Paul JW 3rd, Sun S, Herdy JR, Bieri G, Kramer NJ, Gage FH, Van Den Bosch L, Robberecht W, Gitler AD (2015) Modifiers of C9orf72 dipeptide repeat toxicity connect nucleocytoplasmic transport defects to FTD/ALS. *Nat Neurosci* 18: 1226 – 1229

Ju JS, Fuentealba RA, Miller SE, Jackson E, Piwnica-Worms D, Baloh RH, Weihl CC (2009) Valosin-containing protein (VCP) is required for autophagy and is disrupted in VCP disease. *J Cell Biol* 187: 875 – 888

Kabashi E, Valdmanis PN, Dion P, Spiegelman D, McConkey BJ, Vande Velde C, Bouchard JP, Lacomblez L, Pochigaeva K, Salachas F, Pradat PF, Camu W, Meininger V, Dupre N, Rouleau GA (2008) TARDBP mutations in individuals with sporadic and familial amyotrophic lateral sclerosis. *Nat Genet* 40: 572 – 574

Kabashi E, Lin L, Tradewell ML, Dion PA, Bercier V, Bourgouin P, Rochefort D, Bel Hadj S, Durham HD, Vande Velde C, Rouleau GA, Drapeau P (2010) Gain and loss of function of ALS-related mutations of TARDBP (TDP-43) cause motor deficits in vivo. *Hum Mol Genet* 19: 671 – 683

Kim HJ, Raphael AR, LaDow ES, McGurk L, Weber RA, Trojanowski JQ, Lee VM, Finkbeiner S, Gitler AD, Bonini NM (2014) Therapeutic modulation of eIF2α phosphorylation rescues TDP-43 toxicity in amyotrophic lateral sclerosis disease models. *Nat Genet* 46: 152 – 160

Komatsu M, Waguri S, Chiba T, Murata S, Iwata J, Tanida I, Ueno T, Koike M, Uchiyama Y, Kominami E, Tanaka K (2006) Loss of autophagy in the central nervous system causes neurodegeneration in mice. *Nature* 441: 880 – 884

Koppers M, Blokhuis AM, Westeneng HJ, Terpstra ML, Zundel CA, Vieira de Sá R, Schellevis RD, Waite AJ, Blake DJ, Veldink JH, van den Berg LH, Jeroen Pasterkamp R (2015) C9orf72 ablation in mice does not cause motor neuron degeneration or motor deficits. *Ann Neurol* 78: 426 – 438

Kozlov G, Trempe JF, Khaleghpour K, Kahvejian A, Ekiel I, Gehring K (2001) Structure and function of the C-terminal PABC domain of human poly (A)-binding protein. *Proc Natl Acad Sci U S A* 98: 4409 – 4413

Kwon I, Xiang S, Kato M, Wu L, Theodoropoulos P, Wang T, Kim J, Yun J, Xie Y, McKnight SL (2014) Poly-dipeptides encoded by the C9orf72 repeats bind nucleoli, impede RNA biogenesis, and kill cells. *Science* 345: 1139 – 1145

Lagier-Tourenne C, Baughn M, Rigo F, Sun S, Liu P, Li HR, Jiang J, Watt AT, Chun S, Katz M, Qiu J, Sun Y, Ling SC, Zhu Q, Polymenidou M, Drenner K, Artates JW, McAlonis-Downes M, Markmiller S, Hutt KR *et al* (2013) Targeted degradation of sense and antisense C9orf72 RNA foci as therapy for ALS and frontotemporal degeneration. *Proc Natl Acad Sci U S A* 110: E4530 – E4539

Lattante S, Millecamps S, Stevanin G, Rivaud-Péchoux S, Moigneu C, Camuzat A, Da BS, Mundwiller E, Couarch P, Salachas F, Hannequin D, Meininger V, Pasquier F, Seilhean D, Couratier P, Danel-Brunaud V, Bonnet AM, Tranchant C, LeGuern E, Brice A *et al* (2014) Contribution of ATXN2 intermediary polyQ expansions in a spectrum of neurodegenerative disorders. *Neurology* 83: 990 – 995

Lattante S, Ciura S, Rouleau GA, Kabashi E (2015a) Defining the genetic connection linking amyotrophic lateral sclerosis (ALS) with frontotemporal dementia (FTD). *Trends Genet* 31: 263 – 273

Lattante S, de Calbiac H, Le Ber I, Brice A, Ciura S, Kabashi E (2015b) Sqstm1 knock-down causes a locomotor phenotype ameliorated by rapamycin in a zebrafish model of ALS/FTLD. *Hum Mol Genet* 24: 1682 – 1690

Lazarou M, Sliter DA, Kane LA, Sarraf SA, Wang C, Burman JL, Sideris DP, Fogel AI, Youle RJ (2015) The ubiquitin kinase PINK1 recruits autophagy receptors to induce mitophagy. *Nature* 524: 309 – 314

Lee YB, Chen HJ, Peres JN, Gomez-Deza J, Attig J, Stalekar M, Troakes C, Nishimura AL, Scotter EL, Vance C, Adachi Y, Sardone V, Miller JW, Smith BN, Gallo JM, Ule J, Hirth F, Rogelj B, Houart C, Shaw CE (2013) Hexanucleotide repeats in ALS/FTD form length-dependent RNA foci, sequester RNA binding proteins, and are neurotoxic. *Cell Rep* 5: 1178 – 1186

Levine TP, Daniels RD, Gatta AT, Wong LH, Hayes MJ (2013) The product of C9orf72, a gene strongly implicated in neurodegeneration, is structurally related to DENN Rab-GEFs. *Bioinformatics* 29: 499 – 503

Lomen-Hoerth C, Anderson T, Miller B (2002) The overlap of amyotrophic lateral sclerosis and frontotemporal dementia. *Neurology* 59: 1077 – 1079

Ma X, Helgason E, Phung QT, Quan CL, Iyer RS, Lee MW, Bowman KK, Starovasnik MA, Dueber EC (2012) Molecular basis of Tank-binding kinase 1 activation by transautophosphorylation. *Proc Natl Acad Sci U S A* 109: 9378 – 9383

Majcher V, Goode A, James V, Layfield R (2015) Autophagy receptor defects and ALS-FTLD. *Mol Cell Neurosci* 66 : 43 – 52.

Majounie E, Renton AE, Mok K, Dopper EG, Waite A, Rollinson S, Chiò A, Restagno G, Nicolaou N, Simon-Sanchez J, van Swieten JC, Abramzon Y, Johnson JO, Sendtner M, Pamphlett R, Orrell RW, Mead S, Sidle KC, Houlden H, Rohrer JD *et al* (2012) Frequency of the C9orf72 hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study. *Lancet Neurol* 11: 323 – 330.

Maruyama H, Morino H, Ito H, Izumi Y, Kato H, Watanabe Y, Kinoshita Y, Kamada M, Nodera H, Suzuki H, Komure O, Matsuura S, Kobatake K, Morimoto N, Abe K, Suzuki N, Aoki M, Kawata A, Hirai T, Kato T *et al* (2010) Mutations of optineurin in amyotrophic lateral sclerosis. *Nature* 465: 223 – 226

Matsumoto G, Shimogori T, Hattori N, Nukina N (2015) TBK1 controls autophagosomal engulfment of polyubiquitinated mitochondria through P62/SQSTM1 phosphorylation. *Hum Mol Genet* 24: 4429 – 4442

May S, Hornburg D, Schludi MH, Arzberger T, Rentzsch K, Schwenk BM, Grässer FA, Mori K, Kremmer E, Banzhaf-Strathmann J, Mann M, Meissner F, Edbauer D (2014) C9orf72 FTLD/ALS-associated Gly-Ala dipeptide repeat proteins cause neuronal toxicity and Unc119 sequestration. *Acta Neuropathol* 128: 485 – 503

Mizielinska S, Lashley T, Norona FE, Clayton EL, Ridler CE, Fratta P, Isaacs AM (2013) C9orf72 frontotemporal lobar degeneration is characterised by frequent neuronal sense and antisense RNA foci. *Acta Neuropathol* 126: 845 – 857

Mizielinska S, Grönke S, Niccoli T, Ridler CE, Clayton EL, Devoy A, Moens T, Norona FE, Woollacott IO, Pietrzyk J, Cleverley K, Nicoll AJ, Pickering-Brown S, Dols J, Cabecinha M, Hendrich O, Fratta P, Fisher EM, Partridge L, Isaacs AM (2014) C9orf72 repeat expansions cause neurodegeneration in *Drosophila* through arginine-rich proteins. *Science* 345: 1192 – 1194

Mori K, Lammich S, Mackenzie IR, Forné I, Zilow S, Kretzschmar H, Edbauer D, Janssens J, Kleinberger G, Cruts M, Herms J, Neumann M, Van Broeckhoven C, Arzberger T, Haass C (2013a) hnRNP A3 binds to GGGGCC repeats and is a constituent of P62-positive/TDP43-negative inclusions in the hippocampus of patients with C9orf72 mutations. *Acta Neuropathol* 125: 413 – 423

Mori K, Weng SM, Arzberger T, May S, Rentzsch K, Kremmer E, Schmid B, Kretzschmar HA, Cruts M, Van Broeckhoven C, Haass C, Edbauer D (2013b) The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTLD/ALS. *Science* 339: 1335 – 1338

Neumann M, Sampathu DM, Kwong LK, Truax AC, Micsenyi MC, Chou TT, Bruce J, Schuck T, Grossman M, Clark CM, McCluskey LF, Miller BL, Masliah E, Mackenzie IR, Feldman H, Feiden W, Kretzschmar HA, Trojanowski JQ, Lee VM (2006) Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science* 314: 130 – 133

Nixon RA (2013) The role of autophagy in neurodegenerative disease. *Nat Med* 19: 983 – 997

Nookala RK, Langemeyer L, Pacitto A, Ochoa-Montaño B, Donaldson JC, Blaszczyk BK, Chirgadze DY, Barr FA, Bazan JF, Blundell TL (2012) Crystal structure of folliculin reveals a hidDENN function in genetically inherited renal cancer. *Open Biol* 2: 120071

O'Rourke JG, Bogdanik L, Muhammad AK, Gendron TF, Kim KJ, Austin A, Cady J, Liu EY, Zarrow J, Grant S, Ho R, Bell S, Carmona S, Simpkinson M, Lall D, Wu K, Daughrity L, Dickson DW, Harms MB, Petrucelli L *et al* (2015) C9orf72 BAC Transgenic Mice Display Typical Pathologic Features of ALS/FTD. *Neuron* 88: 892 – 901

Peters OM, Cabrera GT, Tran H, Gendron TF, McKeon JE, Metterville J, Weiss A, Wightman N, Salameh J, Kim J, Sun H, Boylan KB, Dickson D, Kennedy Z, Lin Z, Zhang YJ, Daughrity L, Jung C, Gao FB, Sapp PC *et al* (2015) Human C9ORF72 Hexanucleotide Expansion Reproduces RNA Foci and Dipeptide Repeat Proteins but Not Neurodegeneration in BAC Transgenic Mice. *Neuron* 88: 902 – 909

Petit CS, Roczniak-Ferguson A, Ferguson SM (2013) Recruitment of folliculin to lysosomes supports the amino acid-dependent activation of Rag GTPases. *J Cell Biol* 202: 1107 – 1122

Pilli M, Arko-Mensah J, Ponpuak M, Roberts E, Master S, Mandell MA, Dupont N, Ornatowski W, Jiang S, Bradfute SB, Bruun JA, Hansen TE, Johansen T, Deretic V (2012) TBK-1 promotes autophagy-mediated antimicrobial defense by controlling autophagosome maturation. *Immunity* 37: 223 – 234

Pottier C, Bieniek KF, Finch N, van de Vorst M, Baker M, Perkersen R, Brown P, Ravenscroft T, van Blitterswijk M, Nicholson AM, DeTure M, Knopman DS, Josephs KA, Parisi JE, Petersen RC, Boylan KB, Boeve BF, Graff-Radford NR, Veltman JA, Gilissen C *et al* (2015) Whole-genome sequencing reveals important role for TBK1 and OPTN mutations in frontotemporal lobar degeneration without motor neuron disease. *Acta Neuropathol* 130: 77 – 92

Pulst SM, Nechiporuk A, Nechiporuk T, Gispert S, Chen XN, Lopes-Cendes I, Pearlman S, Starkman S, Orozco-Diaz G, Lunkes A, DeJong P, Rouleau GA, Auburger G, Korenberg JR, Figueroa C, Sahba S (1996) Moderate expansion of a normally biallelic trinucleotide repeat in spinocerebellar ataxia type 2. *Nat Genet* 14: 269 – 276

Renton AE, Majounie E, Waite A, Simón-Sánchez J, Rollinson S, Gibbs JR, Schymick JC, Laaksovirta H, van Swieten JC, Myllykangas L, Kalimo H, Paetau A, Abramzon Y, Remes AM, Kaganovich A, Scholz SW, Duckworth J, Ding J, Harmer DW, Hernandez DG *et al* (2011) A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron* 72: 257 – 268

Ringholz GM, Appel SH, Bradshaw M, Cooke NA, Mosnik DM, Schulz PE (2005) Prevalence and patterns of cognitive impairment in sporadic ALS. *Neurology* 65: 586 – 590

Ross OA, Rutherford NJ, Baker M, Soto-Ortolaza AI, Carrasquillo MM, DeJesus-Hernandez M, Adamson J, Li M, Volkening K, Finger E, Seeley WW, Hatanpaa KJ, Lomen-Hoerth C, Kertesz A, Bigio EH, Lippa C, Woodruff BK, Knopman DS, White CL 3rd, Van Gerpen JA *et al* (2011) Ataxin-2 repeat-length variation and neurodegeneration. *Hum Mol Genet* 20: 3207 – 3212

Rutherford NJ, Zhang YJ, Baker M, Gass JM, Finch NA, Xu YF, Stewart H, Kelley BJ, Kuntz K, Crook RJ, Sreedharan J, Vance C, Sorenson E, Lippa C, Bigio EH, Geschwind DH, Knopman DS, Mitsumoto H, Petersen RC, Cashman NR *et al* (2008) Novel mutations in TARDBP (TDP-43) in patients with familial amyotrophic lateral sclerosis. *PLoS Genet* 4: e1000193

Sanpei K, Takano H, Igarashi S, Sato T, Oyake M, Sasaki H, Wakisaka A, Tashiro K, Ishida Y, Ikeuchi T, Koide R, Saito M, Sato A, Tanaka T, Hanyu S, Takiyama Y, Nishizawa M, Shimizu N, Nomura Y, Segawa M *et al* (1996) Identification of the spinocerebellar ataxia type 2 gene using a direct identification of repeat expansion and cloning technique, DIRECT. *Nat Genet* 14: 277 – 284

Sato T, Iwano T, Kunii M, Matsuda S, Mizuguchi R, Jung Y, Hagiwara H, Yoshihara Y, Yuzaki M, Harada R, Harada A (2014) Rab8a and Rab8b are essential for several apical transport pathways but insufficient for ciliogenesis. *J Cell Sci* 127: 422 – 431

Schaffer BE, Levin RS, Hertz NT, Maures TJ, Schoof ML, Hollstein PE, Benayoun BA, Banko MR, Shaw RJ, Shokat KM, Brunet A (2015) Identification of AMPK Phosphorylation Sites Reveals a Network of Proteins Involved in Cell Invasion and Facilitates Large-Scale Substrate Prediction. *Cell Metab* 22: 907 – 921

Scotter EL, Vance C, Nishimura AL, Lee YB, Chen HJ, Urwin H, Sardone V, Mitchell JC, Rogelj B, Rubinsztein DC, Shaw CE (2014) Differential roles of the ubiquitin proteasome system and autophagy in the clearance of soluble and aggregated TDP-43 species. *J Cell Sci* 127: 1263 – 1278

Seto S, Sugaya K, Tsujimura K, Nagata T, Horii T, Koide Y (2013) Rab39a interacts with phosphatidylinositol 3-kinase and negatively regulates autophagy induced by lipopolysaccharide stimulation in macrophages. *PLoS ONE* 8: e83324

Skibinski G1, Parkinson NJ, Brown JM, Chakrabarti L, Lloyd SL, Hummerich H, Nielsen JE, Hodges JR, Spillantini MG, Thusgaard T, Johannsen P, Sørensen SA, Gydesen S, Fisher EM, Collinge J (2005) Mutations in the endosomal ESCRTIII-complex subunit CHMP2B in frontotemporal dementia. *Nat Genet* 37: 806 – 808.

Sreedharan J, Blair IP, Tripathi VB, Hu X, Vance C, Rogelj B, Ackerley S, Durnall JC, Williams KL, Buratti E, Baralle F, de Belleroche J, Mitchell JD, Leigh PN, Al-Chalabi A, Miller CC, Nicholson G, Shaw CE (2008) TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis. *Science* 319: 1668 – 1672

Sun N, Yun J, Liu J, Malide D, Liu C, Rovira II, Holmström KM, Fergusson MM, Yoo YH, Combs CA, Finkel T (2015) Measuring In Vivo Mitophagy. *Mol Cell* 60: 685 – 696

Suzuki N, Maroof AM, Merkle FT, Koszka K, Intoh A, Armstrong I, Moccia R, Davis-Dusenbery BN, Eggan K (2013) The mouse C9ORF72 ortholog is enriched in neurons known to degenerate in ALS and FTD. *Nat Neurosci* 16: 1725 – 1727

Tao Z, Wang H, Xia Q, Li K, Li K, Jiang X, Xu G, Wang G, Ying Z (2015) Nucleolar stress and impaired stress granule formation contribute to C9orf72 RAN translation-induced cytotoxicity. *Hum Mol Genet* 24: 2426 – 2441

Taylor JP (2015) Multisystem proteinopathy: Intersecting genetics in muscle, bone, and brain degeneration. *Neurology* 85: 658 – 660

Thurston TL, Ryzhakov G, Bloor S, von Muhlinen N, Randow F (2009) The TBK1 adaptor and autophagy receptor NDP52 restricts the proliferation of ubiquitin-coated bacteria. *Nat Immunol* 10: 1215 – 1221

Tsun ZY, Bar-Peled L, Chantranupong L, Zoncu R, Wang T, Kim C, Spooner E, Sabatini DM (2013) The folliculin tumor suppressor is a GAP for the RagC/D GTPases that signal amino acid levels to mTORC1. *Mol Cell* 52: 495 – 505

Uchino A, Takao M, Hatsuta H, Sumikura H, Nakano Y, Nogami A, Saito Y, Arai T, Nishiyama K, Murayama S (2015) Incidence and extent of TDP-43 accumulation in aging human brain. *Acta Neuropathol Commun.* 20: 35

Urushitani M, Sato T, Bamba H, Hisa Y, Tooyama I (2010) Synergistic effect between proteasome and autophagosome in the clearance of polyubiquitinated TDP-43. *J Neurosci Res* 88: 784 – 797

Van Blitterswijk M, van Es MA, Hennekam EA, Dooijes D, van Rheenen W, Medic J, Bourque PR, Schelhaas HJ, van der Kooi AJ, de Visser M, de Bakker PI, Veldink JH, van den Berg LH (2012) Evidence for an oligogenic basis of amyotrophic lateral sclerosis. *Hum Mol Genet* 21: 3776 – 3784

Van Damme P, Veldink JH, van Blitterswijk M, Corveleyn A, van Vught PW, Thijs V, Dubois B, Matthijs G, van den Berg LH, Robberecht W (2011) Expanded ATXN2 CAG repeat size in ALS identifies genetic overlap between ALS and SCA2. *Neurology* 76: 2066 – 2072

Van Deerlin VM, Leverenz JB, Bekris LM, Bird TD, Yuan W, Elman LB, Clay D, Wood EM, Chen-Plotkin AS, Martinez-Lage M, Steinbart E, McCluskey L, Grossman M, Neumann M, Wu IL, Yang WS, Kalb R, Galasko DR, Montine TJ, Trojanowski JQ *et al* (2008) TARDBP mutations in amyotrophic lateral sclerosis with TDP-43 neuropathology: a genetic and histopathological analysis. *Lancet Neurol* 7: 409 – 416

Waite AJ, Bäumer D, East S, Neal J, Morris HR, Ansorge O, Blake DJ (2014) Reduced C9orf72 protein levels in frontal cortex of amyotrophic lateral sclerosis and frontotemporal degeneration brain with the C9ORF72 hexanucleotide repeat expansion. *Neurobiol Aging* 35:1779.e5 – 1779.e13

Wang IF, Guo BS, Liu YC, Wu CC, Yang CH, Tsai KJ, Shen CK (2012) Autophagy activators rescue and alleviate pathogenesis of a mouse model with proteinopathies of the TAR DNA-binding protein 43. *Proc Natl Acad Sci U S A* 109: 15024 – 15029

Wen X, Tan W, Westergard T, Krishnamurthy K, Markandaiah SS, Shi Y, Lin S, Shneider NA, Monaghan J, Pandey UB, Pasinelli P, Ichida JK, Trotti D (2014) Antisense proline-arginine RAN dipeptides linked to C9ORF72-ALS/FTD form toxic nuclear aggregates that initiate in vitro and in vivo neuronal death. *Neuron* 84: 1213 – 1225

West RJ, Lu Y, Marie B, Gao FB, Sweeney ST (2015) Rab8, POSH, and TAK1 regulate synaptic growth in a *Drosophila* model of frontotemporal dementia. *J Cell Biol* 208: 931 – 947

Wild P, Farhan H, McEwan DG, Wagner S, Rogov VV, Brady NR, Richter B, Korac J, Waidmann O, Choudhary C, Dötsch V, Bumann D, Dikic I (2011) Phosphorylation of the autophagy receptor optineurin restricts Salmonella growth. *Science* 333: 228 – 233

Wilke C, Pomper JK, Biskup S, Puskás C, Berg D, Synofzik M (2016) Atypical parkinsonism in C9orf72 expansions: a case report and systematic review of 45 cases from the literature. *J Neurol* 263: 558 – 574

Wilson GR, Sim JC, McLean C, Giannandrea M, Galea CA, Riseley JR, Stephenson SE, Fitzpatrick E, Haas SA, Pope K, Hogan KJ, Gregg RG, Bromhead CJ, Wargowski DS, Lawrence CH, James PA, Churchyard A, Gao Y, Phelan DG, Gillies G *et al* (2014) Mutations in RAB39B cause X-linked intellectual disability and early-onset Parkinson disease with α-synuclein pathology. *Am J Hum Genet* 95: 729 – 735

Wong E, Cuervo AM (2010) Autophagy gone awry in neurodegenerative diseases. *Nat Neurosci* 13: 805 – 811

Xia Q, Wang H, Hao Z, Fu C, Hu Q, Gao F, Ren H, Chen D, Han J, Ying Z, Wang G (2016) TDP-43 loss of function increases TFEB activity and blocks autophagosome-lysosome fusion. *EMBO J* 35: 121 – 142

Xu J, Fotouhi M, McPherson PS (2015) Phosphorylation of the exchange factor DENND3 by ULK in response to starvation activates Rab12 and induces autophagy. *EMBO Rep* 16: 709 – 718

Yokoseki A, Shiga A, Tan CF, Tagawa A, Kaneko H, Koyama A, Eguchi H, Tsujino A, Ikeuchi T, Kakita A, Okamoto K, Nishizawa M, Takahashi H, Onodera O (2008) TDP-43 mutation in familial amyotrophic lateral sclerosis. *Ann Neurol* 63: 538 – 542

Yokoshi M, Li Q, Yamamoto M, Okada H, Suzuki Y, Kawahara Y (2014) Direct binding of Ataxin-2 to distinct elements in 3′ UTRs promotes mRNA stability and protein expression. *Mol Cell* 55: 186 – 198

Zhang D, Iyer LM, He F, Aravind L (2012) Discovery of Novel DENN Proteins: Implications for the Evolution of Eukaryotic Intracellular Membrane Structures and Human Disease. *Front Genet* 13: 283

Zhang YJ, Jansen-West K, Xu YF, Gendron TF, Bieniek KF, Lin WL, Sasaguri H, Caulfield T, Hubbard J, Daughrity L, Chew J, Belzil VV, Prudencio M, Stankowski JN, Castanedes-Casey M, Whitelaw E, Ash PE, DeTure M, Rademakers R, Boylan KB *et al* (2014) Aggregation-prone c9FTD/ALS poly (GA) RAN-translated proteins cause neurotoxicity by inducing ER stress. *Acta Neuropathol* 128: 505 – 524

Zhang K, Donnelly CJ, Haeusler AR, Grima JC, Machamer JB, Steinwald P, Daley EL, Miller SJ, Cunningham KM, Vidensky S, Gupta S, Thomas MA, Hong I, Chiu SL, Huganir RL, Ostrow LW, Matunis MJ, Wang J, Sattler R, Lloyd TE *et al* (2015) The C9orf72 repeat expansion disrupts nucleocytoplasmic transport. *Nature* 525: 56 – 61

Zu T, Liu Y, Bañez-Coronel M, Reid T, Pletnikova O, Lewis J, Miller TM, Harms MB, Falchook AE, Subramony SH, Ostrow LW, Rothstein JD, Troncoso JC, Ranum LP (2013) RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. *Proc Natl Acad Sci U S A* 110: E4968 – E4977

Expanded View Figures

Figure EV1. Validation of the protein interacting with C9ORF72.

A Immunoblot analysis of endogenous proteins found associated by proteomic analysis with control or Flag-HA-tagged C9ORF72 expressed in N2A cells.

B Immunoblot analysis of HA-immunoprecipitated proteins and lysate of HEK293 cells co-expressing HA-tagged long or short splicing variant of C9ORF72 with Flag-tagged SMCR8 and Flag-tagged WDR41.

C Immunoblot analysis of HA-immunoprecipitated proteins and lysate of HEK293 cells co-expressing HA-tagged C9ORF72 and/or HA-tagged SMCR8 and/or HA-tagged WDR41 with Flag-tagged HSC70.

D Left panel, immunoblot analysis of HA-immunoprecipitated proteins and lysate of HEK293 cells co-expressing either HA-tagged C9ORF72, HA-tagged SMCR8, or HA-tagged WDR41 with Flag-tagged RAB8A. Right panel, immunoblot analysis of HA-immunoprecipitated proteins and lysate of HEK293 cells co-expressing either HA-tagged C9ORF72, HA-tagged SMCR8, or HA-tagged WDR41 with Flag-tagged RAB39b.

E Immunoblot analysis of HA-immunoprecipitated proteins and lysate of HEK293 cells co-expressing HA-tagged SMCR8 with various Flag-tagged Rab GTPases.

F Immunoblot analysis of endogenous expression of Smcr8, Rab8a, Rab39b, and Gapdh in mouse adult tissues and primary culture of E18 cortical mouse neurons.

G ^{32}P-radiolabelled GDP release from GST-tagged purified RAB8a, RAB39b, RAB29 (also known as RAB7L1), and RAB32 as a function of increased concentration of recombinant purified HIS-C9ORF72:SMCR8:WDR41 complex purified from baculovirus-infected insect cells. Error bars indicate SEM, $N = 3$.

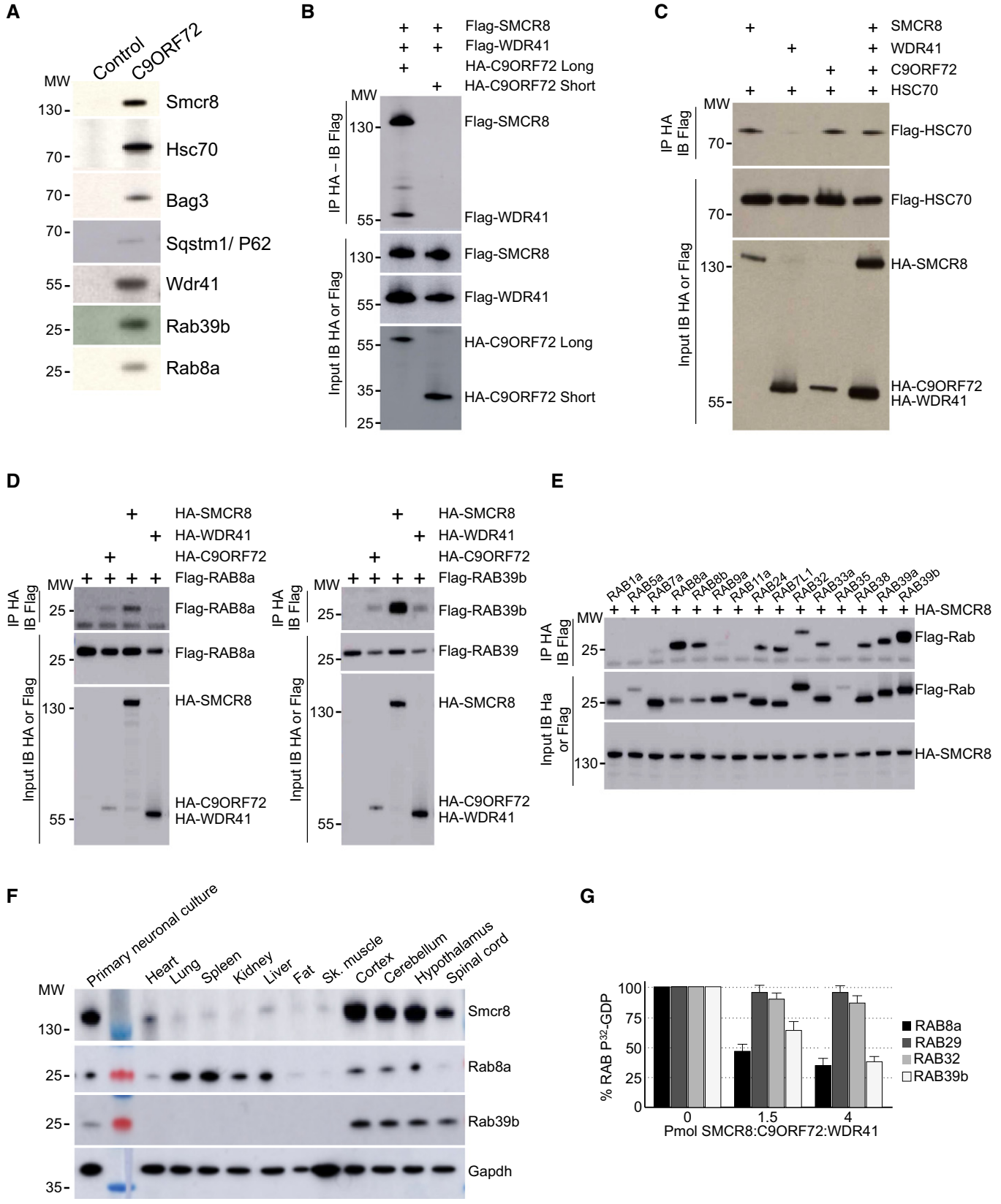


Figure EV1.

Figure EV2. Reduced expression of C9ORF72 partly impairs autophagy.

A Immunoblot analysis of HA-immunoprecipitated proteins and lysate of HEK293 cells co-expressing either HA-tagged C9ORF72, HA-tagged SMCR8, or HA-tagged WDR41 with Flag-tagged P62.

B Immunoblot analysis of HA-immunoprecipitated proteins and lysate of HEK293 cells co-expressing either HA-tagged C9ORF72, HA-tagged SMCR8, or HA-tagged WDR41 with Flag-tagged OPTN.

C Immunoblot analysis of HA-immunoprecipitated proteins and lysate of HEK293 cells co-expressing either HA-tagged RAB8a or HA-tagged RAB39b with Flag-tagged P62.

D Representative images of mouse GT1-7 neuronal cells co-transfected with GFP-RFP-LC3B and either control siRNA or siRNA targeting endogenous *C9orf72* mRNA and treated or not with Torin.

E Upper panel, immunoblot analysis of endogenous LC3B (Map1lc3b), C9orf72 and control Gapdh and actin of GT1-7 neuronal cells transfected with either control siRNA or siRNA targeting *C9orf72* mRNA and treated or not with Torin and/or bafilomycin A. Lower panel, real-time RT–qPCR quantification of endogenous *C9orf72* mRNA expression relative to *Rplp0* mRNA.

F Left panel, representative images of immunofluorescence labeling of endogenous P62 (Sqstm1) on GT1-7 neuronal cells transfected with either control siRNA or siRNA targeting either *C9orf72*, *Smcr8*, or *Wdr41* mRNA. Right panel, quantification of P62 aggregates.

G Left panel, representative merged images of immunofluorescence labeling of transfected constructs (green) and endogenous P62 (Sqstm1, red) on GT1-7 neuronal cells transfected with control siRNA or siRNA targeting *C9orf72* and plasmids expressing control GFP or HA-tagged long or short isoform of C9ORF72. Right panel, quantification of P62 aggregates.

Data information: Scale bars, 10 μm. Nuclei were counterstained with DAPI. Error bars indicate SEM. Student's *t*-test, ***$P < 0.001$, $n = 3$.

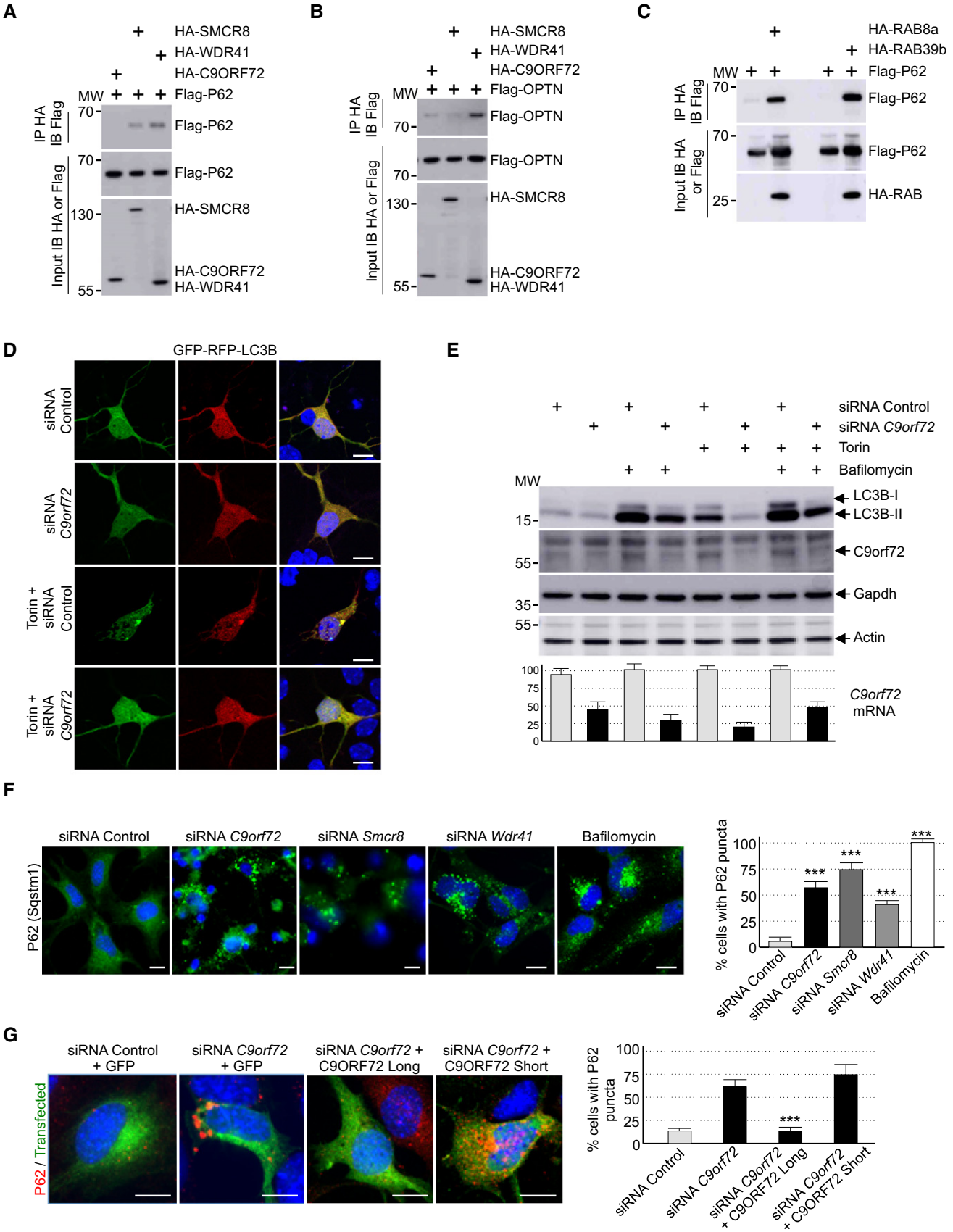


Figure EV2.

Figure EV3. SMCR8 is phosphorylated by ULK1.

A Immunoblot analysis of endogenous Smcr8 or Gapdh in GT1-7 cells transfected with control siRNA or a siRNA targeting the 3′UTR of *Smcr8*.

B Immunoblot analysis of endogenous Tbk1 or Gapdh in GT1-7 cells transfected with control siRNA or a siRNA targeting *Tbk1*.

C Immunoblot analysis of HA-immunoprecipitated proteins and lysate of HEK293 cells co-expressing HA-tagged C9ORF72 and/or HA-tagged SMCR8 and/or HA-tagged WDR41 with Flag-tagged ULK1.

D HA-tagged C9ORF72, HA-tagged SMCR8, and HA-tagged WDR41 were co-expressed in HEK293, immunoprecipitated, and subjected to *in vitro* ULK1 kinase assay in the presence of γ-^{32}P-radiolabelled ATP. Proteins were separated by migration on SDS–PAGE gel and phosphorylation was detected by autoradiography (upper panel), while expression was detected by Western blotting (lower panel).

E Recombinant HIS-tagged C9ORF72:SMCR8:WDR41 complex purified from baculovirus-infected insect cells was phosphorylated *in vitro* by ULK1 and SMCR8 phosphorylation sites were identified by mass spectrometry.

F Upper panel, representative images of immunofluorescence labeling of endogenous P62 (Sqstm1, red) on GT1-7 neuronal cells transfected with control siRNA or siRNA targeting *Tbk1* or *Ulk1* and *Ulk2*. Lower panel, quantification of P62 aggregates.

Data information: Error bars indicate SEM, $n = 3$. Scale bars, 10 μm. Nuclei were counterstained with DAPI.

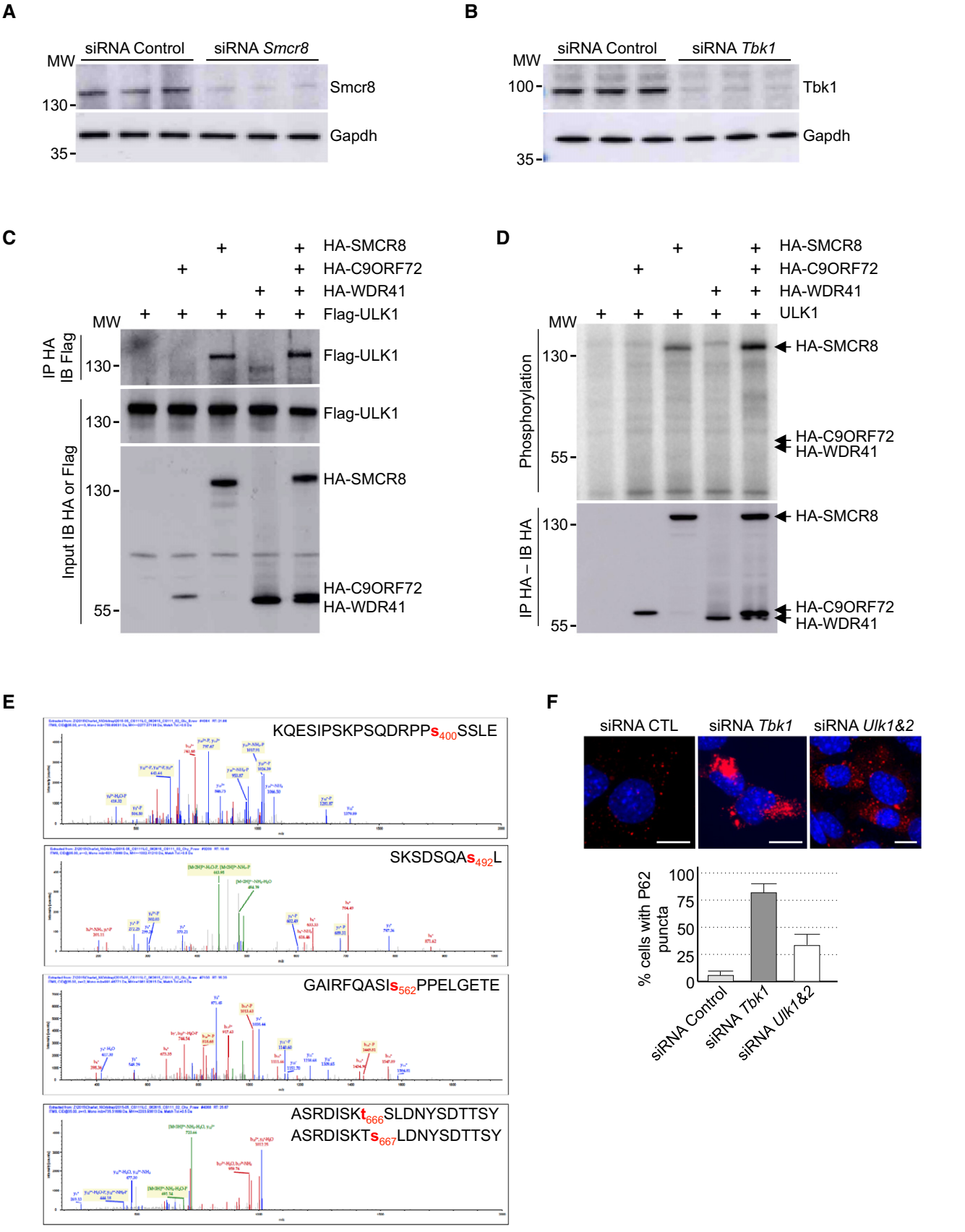


Figure EV3.

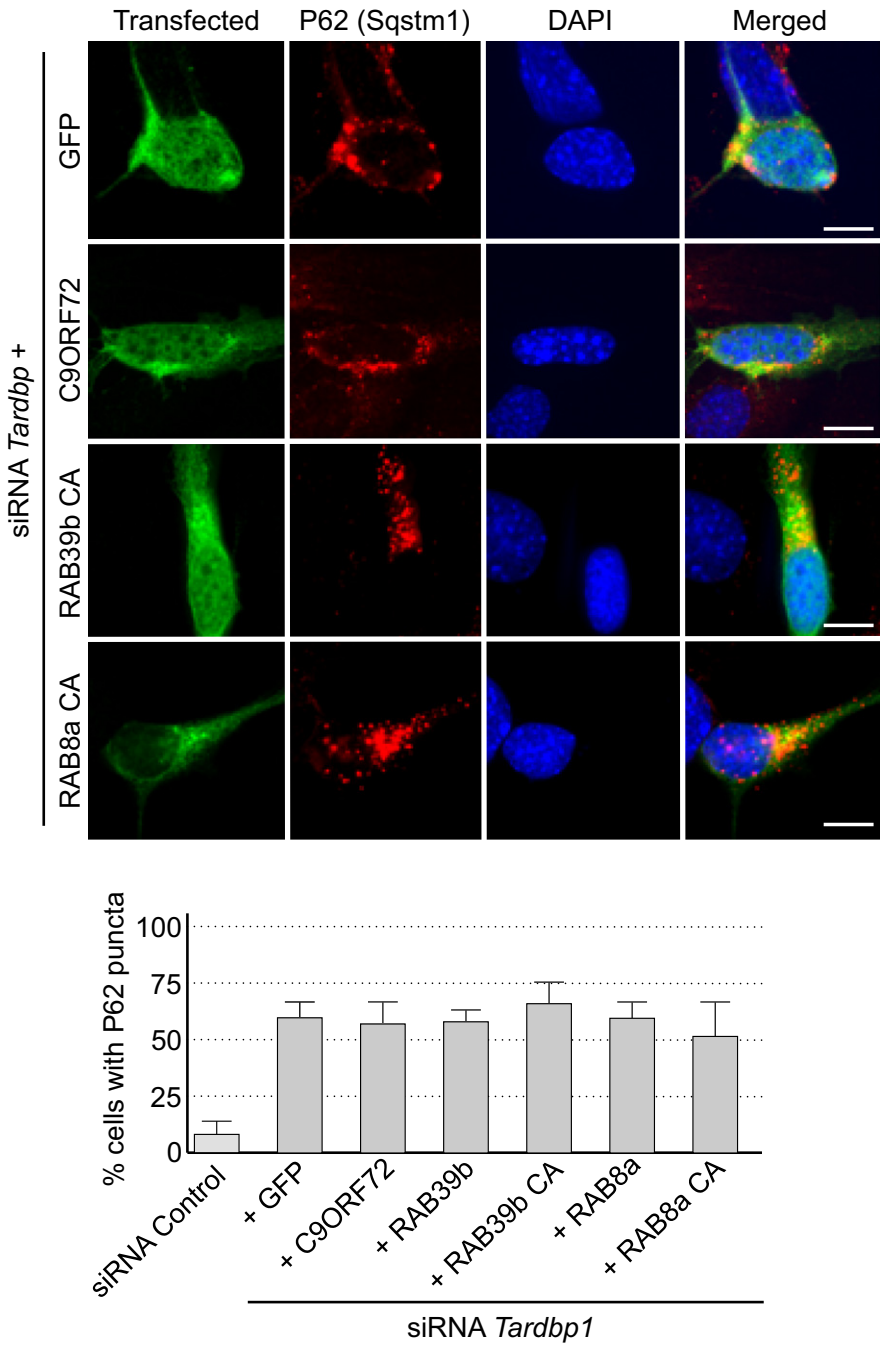


Figure EV4. Reduced expression of C9ORF72 promotes aggregation of TDP-43.
Upper panel, representative images of immunofluorescence labeling of transfected constructs (green) and endogenous P62 (Sqstm1, red) on GT1-7 neuronal cells transfected with siRNA targeting *Tardbp* and plasmids expressing control GFP and HA-tagged C9ORF72 or wild-type or mutant RAB8a or RAB39b. Lower panel, quantification of P62 aggregates. Error bars indicate SEM, $n = 3$.

Figure EV5. Reduced expression of C9ORF72 synergizes Ataxin-2 toxicity.

A Representative images of GT1-7 neuronal cells co-transfected with HA-tagged *ATXN2* with control (Q22x) or intermediate (Q30x) polyQ size and either control siRNA or siRNA targeting endogenous *C9orf72* mRNA.

B, C Immunoblot analysis of endogenous Gapdh or transfected HA-tagged Ataxin-2 with (B) control (Q22x) length or (C) intermediate (Q30x) length of polyQ in GT1-7 cells.

D Representative images of GT1-7 neuronal cells co-transfected with either GFP-tagged mutant SOD1, FUS, HTT, or Ataxin-3 and either control siRNA or siRNA targeting endogenous *C9orf72* mRNA.

E RT–qPCR quantification of endogenous *C9orf72* expression relative to *Gapdh* mRNA in mismatched or *C9orf72* antisense morpholino oligonucleotide (AMO)-injected zebrafish.

F RT–qPCR quantification of exogenous HA-tagged *ATXN2* with normal (Q22x) or intermediate (Q30x) length of polyQ relative to endogenous *Gapdh* mRNA in zebrafish injected with control or HA-tagged ATXN2 constructs and with mismatched AMOs or AMOs against *C9orf72*.

G Tracing of the swimming trajectories of 48 h post-fertilization zebrafish larvae following light touch.

H–J Quantification of the touch-evoked swimming distance (H), average velocity (I), and maximum velocity attained (J) shows no impairment upon the sole injection of HA-tagged *ATXN2* with control (Q22x) or intermediate length of polyQ (Q30x). Similarly, injection of mismatched antisense morpholino oligonucleotides (AMOs) against *C9orf72* alone or with HA-tagged *ATXN2* Q22x or Q30x shows no functional alterations.

Data information: Scale bars, 10 μm. Nuclei were counterstained with DAPI. Error bars indicate SEM, $n = 3$.

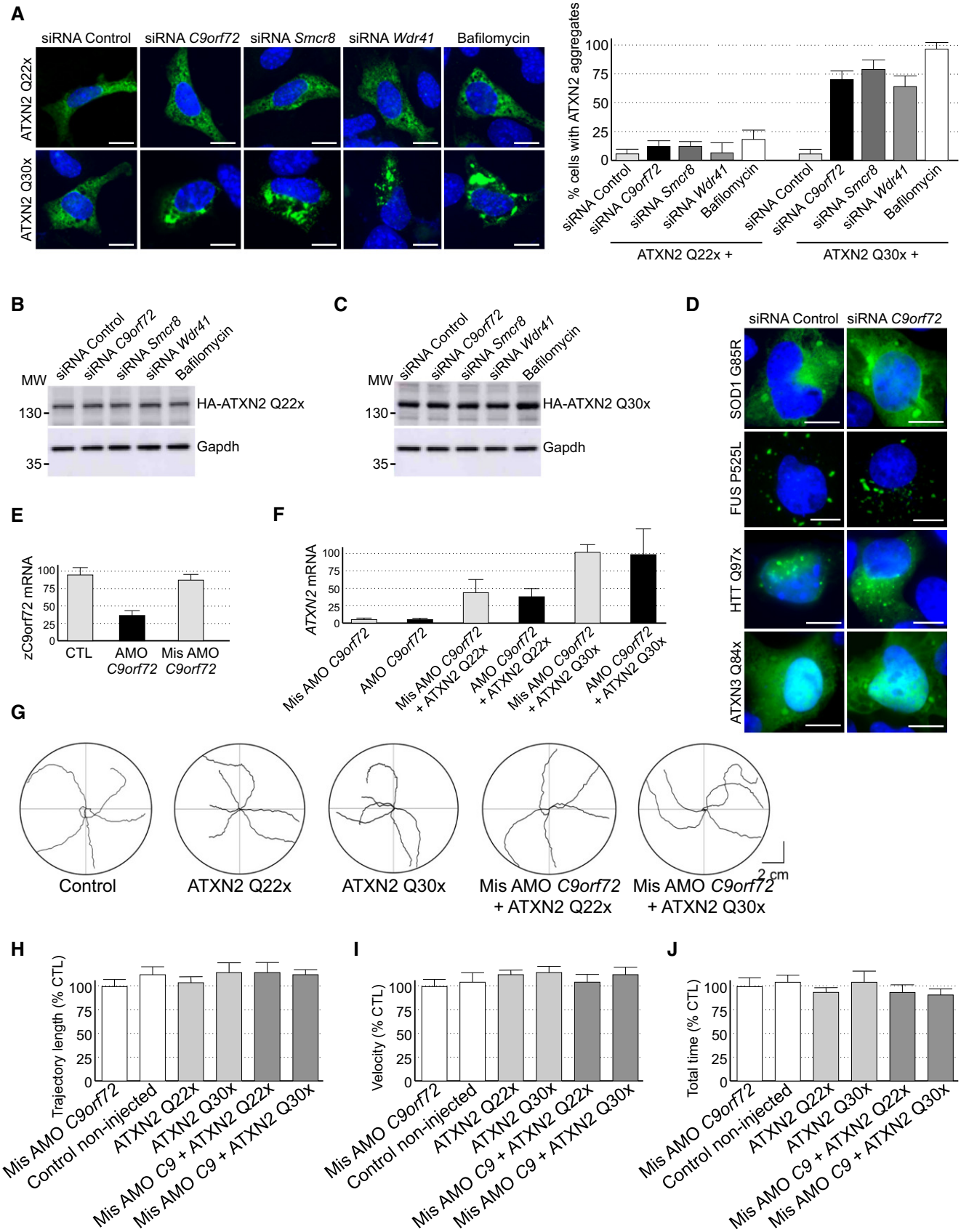

A
siRNA Control
siRNA C9orf72
siRNA Smcr8
siRNA Wdr41
Bafilomycin
ATXN2 Q22x
ATXN2 Q30x
% cells with ATXN2 aggregates
ATXN2 Q22x +
ATXN2 Q30x +
B
MW
130
35
HA-ATXN2 Q22x
Gapdh
C
HA-ATXN2 Q30x
D
SOD1 G85R
FUS P525L
HTT Q97x
ATXN3 Q84x
E
zC9orf72 mRNA
CTL
AMO C9orf72
Mis AMO C9orf72
F
ATXN2 mRNA
Mis AMO C9orf72
AMO C9orf72
Mis AMO C9orf72 + ATXN2 Q22x
AMO C9orf72 + ATXN2 Q22x
Mis AMO C9orf72 + ATXN2 Q30x
AMO C9orf72 + ATXN2 Q30x
G
Control
ATXN2 Q22x
ATXN2 Q30x
Mis AMO C9orf72 + ATXN2 Q22x
Mis AMO C9orf72 + ATXN2 Q30x
2 cm
H
Trajectory length (% CTL)
I
Velocity (% CTL)
J
Total time (% CTL)
Control non-injected
Mis AMO C9 + ATXN2 Q22x
Mis AMO C9 + ATXN2 Q30x

Figure EV5.

Table EV1

Description	Score	Coverage	# Peptides	# PSM	MW [kDa]
Heat shock cognate 71 kDa protein OS=Mus musculus GN=Hspa8 PE=1 SV=1 - [HSP7C_MOUSE]	595	60	31	152	646
Probable helicase senataxin OS=Mus musculus GN=Setx PE=2 SV=1 - [SETX_MOUSE]	211	26	42	54	297
Tyrosine-protein kinase JAK1 OS=Mus musculus GN=Jak1 PE=1 SV=1 - [JAK1_MOUSE]	162	41	32	42	133
Cytoplasmic dynein 1 heavy chain 1 OS=Mus musculus GN=Dync1h1 PE=1 SV=2 - [DYHC1_MOUSE]	156	12	36	41	532
E3 ubiquitin-protein ligase UBR4 OS=Mus musculus GN=Ubr4 PE=1 SV=1 - [UBR4_MOUSE]	102	10	28	28	572
Heat shock protein HSP 90-alpha OS=Mus musculus GN=Hsp90aa1 PE=1 SV=4 - [HS90A_MOUSE]	89	37	19	23	733
BAG family molecular chaperone regulator 3 OS=Mus musculus GN=Bag3 PE=1 SV=2 - [BAG3_MOUSE]	70	31	13	17	577
Smith-Magenis syndrome chromosomal region candidate gene 8 protein homolog OS=Mus musculus GN=Smcr8 PE=1 SV=2 - [SMCR8_MOUSE]	45	26	11	12	105
WD repeat-containing protein 41 OS=Mus musculus GN=Wdr41 PE=2 SV=1 - [WDR41_MOUSE]	28	24	7	8	51
Ras GTPase-activating-like protein IQGAP1 OS=Mus musculus GN=Iqgap1 PE=1 SV=2 - [IQGA1_MOUSE]	33	6	7	8	189
Vesicle-fusing ATPase OS=Mus musculus GN=Nsf PE=1 SV=2 - [NSF_MOUSE]	26	14	7	7	83
Activating molecule in BECN1-regulated autophagy protein 1 OS=Mus musculus GN=Ambra1 PE=1 SV=1 - [AMRA1_MOUSE]	31	9	6	7	143
Sequestosome-1 OS=Mus musculus GN=Sqstm1 PE=1 SV=1 - [SQSTM_MOUSE]	29	31	6	6	48
TAR DNA-binding protein 43 OS=Mus musculus GN=Tardbp PE=1 SV=1 - [TADBP_MOUSE]	23	22	6	6	45
Dynactin subunit 1 OS=Mus musculus GN=Dctn1 PE=1 SV=3 - [DCTN1_MOUSE]	22	8	6	6	142
STIP1 homology and U box-containing protein 1 OS=Mus musculus GN=Stub1 PE=1 SV=1 - [CHIP_MOUSE]	22	3	5	6	304
Ras-related protein Rab-39B OS=Mus musculus GN=Rab39b PE=2 SV=1 - [RB39B_MOUSE]	17	33	5	5	25
RB1-inducible coiled-coil protein 1 OS=Mus musculus GN=Rb1cc1 PE=1 SV=3 - [RBCC1_MOUSE]	12	3	4	4	182
Ras-related protein Rab-8A OS=Mus musculus GN=Rab8a PE=1 SV=2 - [RAB8A_MOUSE]	12	18	3	3	24
Vesicle-trafficking protein SEC22b OS=Mus musculus GN=Sec22b PE=1 SV=3 - [SC22B_MOUSE]	12	17	3	3	25
Regulatory-associated protein of mTOR OS=Mus musculus GN=Rptor PE=1 SV=1 - [RPTOR_MOUSE]	11	3	3	3	149
Vacuolar protein sorting-associated protein 16 homolog OS=Mus musculus GN=Vps16 PE=1 SV=3 - [VPS16_MOUSE]	10	5	3	3	95
Ras-related protein Rab-33B OS=Mus musculus GN=Rab33b PE=1 SV=1 - [RB33B_MOUSE]	9	3	3	3	26
26S proteasome non-ATPase regulatory subunit 8 OS=Mus musculus GN=Psmd8 PE=1 SV=2 - [PSMD8_MOUSE]	8	13	2	2	40
Proteasome subunit alpha type-4 OS=Mus musculus GN=Psma4 PE=1 SV=1 - [PSA4_MOUSE]	7	13	2	2	29
Ras-related protein Rab-26 OS=Mus musculus GN=Rab26 PE=2 SV=2 - [RAB26_MOUSE]	7	13	2	2	22
26S proteasome non-ATPase regulatory subunit 10 OS=Mus musculus GN=Psmd10 PE=1 SV=3 - [PSD10_MOUSE]	8	13	2	2	25
Syntaxin-17 OS=Mus musculus GN=Stx17 PE=1 SV=1 - [STX17_MOUSE]	7	11	2	2	33
E3 ubiquitin-protein ligase RNF126 OS=Mus musculus GN=Rnf126 PE=1 SV=1 - [RN126_MOUSE]	8	9	2	2	34
Serine/threonine-protein kinase TBK1 OS=Mus musculus GN=Tbk1 PE=1 SV=1 - [TBK1_MOUSE]	8	4	2	2	83
Ras-related protein Rab-6A OS=Mus musculus GN=Rab6a PE=1 SV=4 - [RAB6A_MOUSE]	7	7	2	2	24
Vacuolar protein sorting-associated protein 4B OS=Mus musculus GN=Vps4b PE=1 SV=2 - [VPS4B_MOUSE]	8	7	2	2	49
Sorting nexin-6 OS=Mus musculus GN=Snx6 PE=1 SV=2 - [SNX6_MOUSE]	7	7	2	2	47
26S proteasome non-ATPase regulatory subunit 4 OS=Mus musculus GN=Psmd4 PE=1 SV=1 - [PSMD4_MOUSE]	8	7	2	2	41
Dynactin subunit 2 OS=Mus musculus GN=Dctn2 PE=1 SV=3 - [DCTN2_MOUSE]	7	6	2	2	44
TBC1 domain family member 22A OS=Mus musculus GN=Tbc1d22a PE=1 SV=3 - [TB22A_MOUSE]	8	6	2	2	59
Sorting and assembly machinery component 50 homolog OS=Mus musculus GN=Samm50 PE=1 SV=1 - [SAM50_MOUSE]	7	5	2	2	52
E3 ubiquitin-protein ligase RLIM OS=Mus musculus GN=Rlim PE=1 SV=2 - [RNF12_MOUSE]	10	5	2	2	66
Sorting nexin-8 OS=Mus musculus GN=Snx8 PE=2 SV=1 - [SNX8_MOUSE]	7	5	2	2	52
Atlastin-2 OS=Mus musculus GN=Atl2 PE=1 SV=1 - [ATLA2_MOUSE]	7	5	2	2	66
Synaptic vesicle glycoprotein 2A OS=Mus musculus GN=Sv2a PE=1 SV=1 - [SV2A_MOUSE]	7	5	2	2	83
WD repeat-containing protein 35 OS=Mus musculus GN=Wdr35 PE=2 SV=3 - [WDR35_MOUSE]	6	3	2	2	134
5'-AMP-activated protein kinase catalytic subunit alpha-1 OS=Mus musculus GN=Prkaa1 PE=1 SV=2 - [AAPK1_MOUSE]	7	3	2	2	64
E3 ubiquitin-protein ligase HUWE1 OS=Mus musculus GN=Huwe1 PE=1 SV=5 - [HUWE1_MOUSE]	7	1	2	2	482

Table EV1. Proteins potentially interacting with C9ORF72.
Proteins associated with tandem-tagged C9ORF72 expressed in mouse N2A neuronal cells were captured through consecutive anti-Flag and anti-HA affinity purification steps and identified by orbitrap ion trap mass analyzer.

IV- CONCLUSION GENERALE

Durant mes dix années de recherche sur les maladies neurodégénératives dues à des expansions de répétitions dans les régions « non codantes » du génome, j'ai participé à deux découvertes qui permettront peut être de mieux comprendre les mécanismes moléculaires sous-jacents aux pathologies FXTAS et SLA/DFT.

Notamment, mon travail sur FXTAS montre que des expansions limitées (55 à 200) de répétitions CGG peuvent être traduites par l'utilisation d'un codon d'initiation non-canonique de type ACG. Le développement de cellules iPS de patients et de modèles murins montre que cette traduction conduit à l'expression d'une nouvelle protéine, FMRpolyG, toxique pour l'architecture nucléaire des neurones. Nous pouvons espérer que cette protéine puisse constituer un biomarqueur ou une cible thérapeutique pour FXTAS. De plus, ce travail permettra peut être de mieux comprendre comment des expansions de répétitions peuvent être traduites en absence de codon d'initiation ATG.

Concernant la SLA/DFT due à une expansion de répétitions G_4C_2 dans le gène *C9ORF72*, j'ai montré que la protéine C9ORF72 forme un complexe impliqué dans la régulation de l'autophagie. Ces résultats ont été rapidement confirmés depuis par plusieurs études, mais il reste aujourd'hui à comprendre comment et si une diminution de l'expression de *C9ORF72* contribue à la pathologie de la SLA/DFT. En effet, aucune mutation perte de fonction classique n'a été identifiée dans les patients SLA/DFT et les souris knockout pour *C9Orf72* ne montrent pas de signes de neurodégénérescence. Toutefois, l'observation d'interactions moléculaires et cellulaires entre TBK1, C9ORF72, OPTN et p62/SQSTM1 qui sont tous impliqués dans l'autophagie et sont retrouvés mutés chez les patients SLA/DFT suggère un mécanisme commun altéré dans la SLA/DFT. Il reste donc à déterminer si la diminution d'expression de *C9ORF72* pourrait sensibiliser les neurones à un second stress. Ce stress pouvant être l'expression de DPR traduits depuis les répétitions G_4C_2.

En conclusion, nous pouvons espérer qu'une meilleure compréhension des mécanismes moléculaires impliqués dans les maladies neurodégénératives FXTAS et SLA/DFT permettra de proposer de nouvelles approches thérapeutiques pour ces maladies.

BIBLIOGRAPHIE

-A-

Al-Chalabi A, Hardiman O. **(2013)** The epidemiology of ALS: a conspiracy of genes, environment and time. *Nat Rev Neurol.* 9(11):617-28.

Al-Sarraj S, King A, Troakes C, Smith B, Maekawa S, Bodi I, Rogelj B, Al-Chalabi A, Hortobágyi T, Shaw CE. **(2011)** p62 positive, TDP-43 negative, neuronal cytoplasmic and intranuclear inclusions in the cerebellum and hippocampus define the pathology of C9orf72-linked FTLD and MND/ALS. *Acta Neuropathol*. 122(6):691-702.

Almeida S, Gascon E, Tran H, Chou HJ, Gendron TF, Degroot S, Tapper AR, Sellier C, Charlet-Berguerand N, Karydas A, Seeley WW, Boxer AL, Petrucelli L, Miller BL, Gao FB. **(2013)** Modeling key pathological features of frontotemporal dementia with C9ORF72 repeat expansion in iPSC-derived human neurons. *Acta Neuropathol.* 126(3):385-99.

Amick J, Roczniak-Ferguson A, Ferguson SM. **(2016)** C9orf72 binds SMCR8, localizes to lysosomes, and regulates mTORC1 signaling. *Mol Biol Cell.* 27(20):3040-3051.

Arai T, Hasegawa M, Akiyama H, Ikeda K, Nonaka T, Mori H, Mann D, Tsuchiya K, Yoshida M, Hashizume Y, Oda T. **(2006)** TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Biochem Biophys Res Commun.* 351(3):602-11.

Arighi A, Fumagalli GG, Jacini F, Fenoglio C, Ghezzi L, Pietroboni AM, De Riz M, Serpente M, Ridolfi E, Bonsi R, Bresolin N, Scarpini E, Galimberti D. **(2012)** Early onset behavioral variant frontotemporal dementia due to the C9ORF72 hexanucleotide repeat expansion: psychiatric clinical presentations. *J Alzheimers Dis.* 31(2):447-52.

Arocena DG, Iwahashi CK, Won N, Beilina A, Ludwig AL, Tassone F, Schwartz PH, Hagerman PJ. **(2005)** Induction of inclusion formation and disruption of lamin A/C structure by premutation CGG-repeat RNA in human cultured neural cells. *Hum Mol Genet.* 14(23):3661-71.

Ash PE, Bieniek K.F., Gendron T.F., Caulfield T., Lin W.L., Dejesus-Hernandez M., van Blitterswijk M.M., Jansen-West K., Paul J.W., Rademakers R., Boylan K.B., Dickson D.W., Petrucelli L. **(2013)** Unconventional translation of C9ORF72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. *Neuron.* 77 ; 639-646.

Atanasio A, Decman V, White D, Ramos M, Ikiz B, Lee HC, Siao CJ, Brydges S, LaRosa E, Bai Y, Fury W, Burfeind P, Zamfirova R, Warshaw G, Orengo J, Oyejide A, Fralish M, Auerbach W, Poueymirou W, Freudenberg J, Gong G, Zambrowicz B, Valenzuela D, Yancopoulos G, Murphy A, Thurston G, Lai KM. **(2016)** C9orf72 ablation causes immune dysregulation characterized by leukocyte expansion, autoantibody production, and glomerulonephropathy in mice. *Sci Rep.* 6:23204.

-B-

Baborie A, Griffiths TD, Jaros E, Perry R, McKeith IG, Burn DJ, Masuda-Suzukake M, Hasegawa M, Rollinson S, Pickering-Brown S, Robinson AC, Davidson YS, Mann DM. **(2015)** Accumulation of dipeptide repeat proteins predates that of TDP-43 in frontotemporal lobar degeneration associated with hexanucleotide repeat expansions in C9ORF72 gene. *Neuropathol Appl Neurobiol.* 41(5):601-12.

Bacalman S, Farzin F, Bourgeois JA, Cogswell J, Goodlin-Jones BL, Gane LW, Grigsby J, Leehey MA, Tassone F, Hagerman RJ. **(2006)** Psychiatric phenotype of the fragile X-associated tremor/ataxia syndrome (**FXTAS**) in males: newly described fronto-subcortical dementia. *J Clin Psychiatry.* 67(1):87-94.

Baker M, Mackenzie IR, Pickering-Brown SM, Gass J, Rademakers R, Lindholm C, Snowden J, Adamson J, Sadovnick AD, Rollinson S, Cannon A, Dwosh E, Neary D, Melquist S, Richardson A, Dickson D, Berger Z, Eriksen J, Robinson T, Zehr C, Dickey CA, Crook R, McGowan E, Mann D, Boeve B, Feldman H, Hutton M.**(2006)** Mutations in progranulin cause tau-negative frontotemporal dementia linked to chromosome 17. *Nature.* 442(7105):916-9. Epub 2006 Jul 16.

Bañez-Coronel M, Ayhan F, Tarabochia AD, Zu T, Perez BA, Tusi SK, Pletnikova O, Borchelt DR, Ross CA, Margolis RL, Yachnis AT, Troncoso JC, Ranum LP. **(2015)** RAN Translation in Huntington Disease. *Neuron*. 88(4):667-77.

Bannwarth S, Ait-El-Mkadem S, Chaussenot A, Genin EC, Lacas-Gervais S, Fragaki K, Berg-Alonso L, Kageyama Y, Serre V, Moore D, Verschueren A, Rouzier C, Le Ber I, Augé G, Cochaud C, Lespinasse F, N'Guyen K, de Septenville A, Brice A, Yu-Wai-Man P, Sesaki H, Pouget J, Paquis-Flucklinger V. **(2014)** Two novel mutations in conserved codons indicate that CHCHD10 is a gene associated with motor neuron disease. *Brain.* 137(Pt 12):e310.

Barmada SJ, Serio A, Arjun A, Bilican B, Daub A, Ando DM, Tsvetkov A, Pleiss M, Li X, Peisach D, Shaw C, Chandran S, Finkbeiner S. **(2014)** Autophagy induction enhances TDP43 turnover and survival in neuronal ALS models. *Nat Chem Biol.* 10(8):677-85.

Bar-Peled L, Chantranupong L, Cherniack AD, Chen WW, Ottina KA, Grabiner BC, Spear ED, Carter SL, Meyerson M, Sabatini DM. **(2013)** A Tumor suppressor complex with GAP activity for the Rag GTPases that signal amino acid sufficiency to mTORC1. *Science.* 340(6136):1100-6.

Battistella G, Niederhauser J, Fornari E, Hippolyte L, Gronchi Perrin A, Lesca G, Forzano F, Hagmann P, Vingerhoets FJ, Draganski B, Maeder P, Jacquemont S. **(2013)** Brain structure in asymptomatic FMR1 premutation carriers at risk for fragile X-associated tremor/ataxia syndrome. *Neurobiol Aging.*34(6):1700-7.

Baumann T, Tolnayb M, Monschc A **(2009)** Aphasie primaire progressive: mémoire sans parole. *Forum Med Suisse.* 9(37):646.

Behrends C, Sowa ME, Gygi SP, Harper JW. **(2010)** Network organization of the human autophagy system. *Nature.* 466(7302):68-76.

Belzil VV, Bauer PO, Prudencio M, Gendron TF, Stetler CT, Yan IK, Pregent L, Daughrity L, Baker MC, Rademakers R, Boylan K, Patel TC, Dickson DW, Petrucelli L. **(2013)** Reduced C9orf72 gene expression in c9FTD/ALS is caused by histone trimethylation, an epigenetic event detectable in blood. *Acta Neuropathol.* 126(6):895-905.

Belzil VV, Bauer PO, Gendron TF, Murray ME, Dickson D, Petrucelli L. (2014) Characterization of DNA hypermethylation in the cerebellum of c9FTD/ALS patients. *Brain Res.* 1584:15-21.

Benussi L, Rossi G, Glionna M, Tonoli E, Piccoli E, Fostinelli S, Paterlini A, Flocco R, Albani D, Pantieri R, Cereda C, Forloni G, Tagliavini F, Binetti G, Ghidoni R. **(2014)** C9ORF72 hexanucleotide repeat number in frontotemporal lobar degeneration: a genotype-phenotype correlation study. *J Alzheimers Dis.* 2014;38(4):799-808.

Boeve BF, Graff-Radford NR. **(2012)** Cognitive and behavioral features of c9FTD/ALS. *Alzheimers Res Ther.* 4(4):29.

Boeynaems S, Bogaert E, Michiels E, Gijselinck I, Sieben A, Jovičić A, De Baets G, Scheveneels W, Steyaert J, Cuijt I, Verstrepen KJ, Callaerts P, Rousseau F, Schymkowitz J, Cruts M, Van Broeckhoven C, Van Damme P, Gitler AD, Robberecht W, Van Den Bosch L. **(2016)** Drosophila screen connects nuclear transport genes to DPR pathology in c9ALS/FTD. *Sci Rep.* 6:20877.

Brouwer JR, Willemsen R, Oostra BA. **(2009)** Microsatellite repeat instability and neurological disease. *Bioessays.* 31(1):71-83.

Brunberg JA, Jacquemont S, Hagerman RJ, Berry-Kravis EM, Grigsby J, Leehey MA, Tassone F, Brown WT, Greco CM, Hagerman PJ. **(2002)** Fragile X premutation carriers: characteristic MR imaging findings of adult male patients with progressive cerebellar and cognitive dysfunction. *AJNR Am J Neuroradiol.* 23(10):1757-66.

Buijsen RA, Sellier C, Severijnen LA, Oulad-Abdelghani M, Verhagen RF, Berman RF, Charlet-Berguerand N, Willemsen R, Hukema RK. **(2014)** FMRpolyG-positive inclusions in CNS and non-CNS organs of a fragile X premutation carrier with fragile X-associated tremor/ataxia syndrome. *Acta Neuropathol Commun*. 2:162.

Burberry A, Suzuki N, Wang JY, Moccia R, Mordes DA, Stewart MH, Suzuki-Uematsu S, Ghosh S, Singh A, Merkle FT, Koszka K, Li QZ, Zon L, Rossi DJ, Trowbridge JJ, Notarangelo LD, Eggan K. **(2016)** Loss-of-function mutations in the C9ORF72 mouse ortholog cause fatal autoimmune disease. *Sci Transl Med.* 8(347):347ra93.

Burrell JR, Kiernan MC, Vucic S, Hodges JR. (**2011)** Motor neuron dysfunction in frontotemporal dementia. *Brain.* 134(Pt 9):2582-94.

-C-

Chew J, Gendron TF, Prudencio M, Sasaguri H, Zhang YJ, Castanedes-Casey M, Lee CW, Jansen-West K, Kurti A, Murray ME, Bieniek KF, Bauer PO, Whitelaw EC, Rousseau L, Stankowski JN, Stetler C, Daughrity LM, Perkerson EA, Desaro P, Johnston A, Overstreet K, Edbauer D, Rademakers R, Boylan KB, Dickson DW, Fryer JD, Petrucelli L. **(2015)** Neurodegeneration. C9ORF72 repeat expansions in mice cause TDP-43 pathology, neuronal loss, and behavioral deficits. *Science.* 348(6239):1151-4.

Cirulli ET, Lasseigne BN, Petrovski S, Sapp PC, Dion PA, Leblond CS, Couthouis J, Lu YF, Wang Q, Krueger BJ, Ren Z, Keebler J, Han Y, Levy SE, Boone BE, Wimbish JR, Waite LL, Jones AL, Carulli JP, Day-Williams AG, Staropoli JF, Xin WW, Chesi A, Raphael AR, McKenna-Yasek D, Cady J, Vianney de Jong JM, Kenna KP, Smith BN, Topp S, Miller J, Gkazi A; FALS Sequencing Consortium, Al-Chalabi A, van den Berg LH, Veldink J, Silani V, Ticozzi N, Shaw CE, Baloh RH, Appel S, Simpson E, Lagier-Tourenne C, Pulst SM, Gibson S, Trojanowski JQ, Elman L, McCluskey L, Grossman M, Shneider NA, Chung WK, Ravits JM, Glass JD, Sims KB, Van Deerlin VM, Maniatis T, Hayes SD, Ordureau A, Swarup S, Landers J, Baas F, Allen AS, Bedlack RS, Harper JW, Gitler AD, Rouleau GA, Brown R, Harms MB, Cooper GM, Harris T, Myers RM, Goldstein DB. **(2015)** Exome sequencing in amyotrophic lateral sclerosis identifies risk genes and pathways. *Science.* 347(6229):1436-41.

Ciura S, Lattante S, Le Ber I, Latouche M, Tostivint H, Brice A, Kabashi E. **(2013)** Loss of function of C9orf72 causes motor deficits in a zebrafish model of amyotrophic lateral sclerosis. *Ann Neurol*. 74(2):180-7.

Clark LN, Poorkaj P, Wszolek Z, Geschwind DH, Nasreddine ZS, Miller B, Li D, Payami H, Awert F, Markopoulou K, Andreadis A, D'Souza I, Lee VM, Reed L, Trojanowski JQ, Zhukareva V, Bird T, Schellenberg G, Wilhelmsen KC. **(1998)** Pathogenic implications of mutations in the tau gene in pallido-ponto-nigral degeneration and related neurodegenerative disorders linked to chromosome 17. *Proc Natl Acad Sci U S A.* 95(22):13103-7.

Cohen S, Masyn K, Adams J, Hessl D, Rivera S, Tassone F, Brunberg J, DeCarli C, Zhang L, Cogswell J, Loesch D, Leehey M, Grigsby J, Hagerman PJ, Hagerman R. **(2006)** Molecular and imaging correlates of the fragile X-associated tremor/ataxia syndrome. *Neurology.* 67(8):1426-31.

Conlon EG, Lu L, Sharma A, Yamazaki T, Tang T, Shneider NA, Manley JL. **(2016)** The C9ORF72 GGGGCC expansion forms RNA G-quadruplex inclusions and sequesters hnRNP H to disrupt splicing in ALS brains. *Elife.* 5. pii: e17820.

Cooper-Knock J, Walsh MJ, Higginbottom A, Robin Highley J, Dickman MJ, Edbauer D, Ince PG, Wharton SB, Wilson SA, Kirby J, Hautbergue GM, Shaw PJ. **(2014)** Sequestration of multiple RNA recognition motif-containing proteins by C9orf72 repeat expansions. *Brain.* 137(Pt 7):2040-51.

Coppola G, Chinnathambi S, Lee JJ, Dombroski BA, Baker MC, Soto-Ortolaza AI, Lee SE, Klein E, Huang AY, Sears R, Lane JR, Karydas AM, Kenet RO, Biernat J, Wang LS, Cotman CW, Decarli CS, Levey AI, Ringman JM, Mendez MF, Chui HC, Le Ber I, Brice A, Lupton MK, Preza E, Lovestone S, Powell J, Graff-Radford N, Petersen RC, Boeve BF, Lippa CF, Bigio EH, Mackenzie I, Finger E, Kertesz A, Caselli RJ, Gearing M, Juncos JL, Ghetti B, Spina S, Bordelon YM, Tourtellotte WW, Frosch MP, Vonsattel JP, Zarow C, Beach TG, Albin RL, Lieberman AP, Lee VM, Trojanowski JQ, Van Deerlin VM, Bird TD, Galasko DR, Masliah E, White CL, Troncoso JC, Hannequin D, Boxer AL, Geschwind MD, Kumar S,

Mandelkow EM, Wszolek ZK, Uitti RJ, Dickson DW, Haines JL, Mayeux R, Pericak-Vance MA, Farrer LA; Alzheimer's Disease Genetics Consortium, Ross OA, Rademakers R, Schellenberg GD, Miller BL, Mandelkow E, Geschwind DH. **(2012)** Evidence for a role of the rare p.A152T variant in MAPT in increasing the risk for FTD-spectrum and Alzheimer's diseases. *Hum Mol Genet.*21(15):3500-12.

Couratier P, Druet-Cabanac M, Truong CT, Bernet-Bernady P, Dumas M, Vallat JM, Preux PM. **(2000)** Interest of a computerized ALS database in the diagnosis and follow-up of patients with ALS. *Rev Neurol (Paris)*. 156(4):357-63.

Cruts M, Gijselinck I, van der Zee J, Engelborghs S, Wils H, Pirici D, Rademakers R, Vandenberghe R, Dermaut B, Martin JJ, van Duijn C, Peeters K, Sciot R, Santens P, De Pooter T, Mattheijssens M, Van den Broeck M, Cuijt I, Vennekens K, De Deyn PP, Kumar-Singh S, Van Broeckhoven C. **(2006)** Null mutations in progranulin cause ubiquitin-positive frontotemporal dementia linked to chromosome 17q21. *Nature.* 442(7105):920-4.

-D-

Damrath E, Heck MV, Gispert S, Azizov M, Nowock J, Seifried C, Rüb U, Walter M, Auburger G. **(2012)** ATXN2-CAG42 sequesters PABPC1 into insolubility and induces FBXW8 in cerebellum of old ataxic knock-in mice. *PLoS Genet.* 2012;8(8):e1002920.

Davidson YS, Barker H, Robinson AC, Thompson JC, Harris J, Troakes C, Smith B, Al-Saraj S, Shaw C, Rollinson S, Masuda-Suzukake M, Hasegawa M, Pickering-Brown S, Snowden JS, Mann DM. **(2014)** Brain distribution of dipeptide repeat proteins in frontotemporal lobar degeneration and motor neurone disease associated with expansions in C9ORF72. *Acta Neuropathol Commun.* 2:70.

Davidson Y, Robinson AC, Liu X, Wu D, Troakes C, Rollinson S, Masuda-Suzukake M, Suzuki G, Nonaka T, Shi J, Tian J, Hamdalla H, Ealing J, Richardson A, Jones M, Pickering-Brown S, Snowden JS, Hasegawa M, Mann DM. **(2016)** Neurodegeneration in frontotemporal lobar degeneration and motor neurone disease associated with expansions in C9orf72 is linked to TDP-43 pathology and not associated with aggregated forms of dipeptide repeat proteins. *Neuropathol Appl Neurobiol.* 42(3):242-54.

de la Luna S, Allen KE, Mason SL, La Thangue NB. (1999) Integration of a growth-suppressing BTB/POZ domain protein with the DP component of the E2F transcription factor. *EMBO J.* 18(1):212-28.

Dechat T, Korbei B, Vaughan OA, Vlcek S, Hutchison CJ, Foisner R. **(2000)** Lamina-associated polypeptide 2alpha binds intranuclear A-type lamins. *J Cell Sci.* 113 Pt 19:3473-84.

DeJesus-Hernandez M, Mackenzie IR, Boeve BF, Boxer AL, Baker M, Rutherford NJ, Nicholson AM, Finch NA, Flynn H, Adamson J, Kouri N, Wojtas A, Sengdy P, Hsiung GY, Karydas A, Seeley WW, Josephs KA, Coppola G, Geschwind DH, Wszolek ZK, Feldman H, Knopman DS, Petersen RC, Miller BL, Dickson DW, Boylan KB, Graff-Radford NR, Rademakers R. **(2011)** Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron.* 72(2):245-56.

Del Bo R, Tiloca C, Pensato V, Corrado L, Ratti A, Ticozzi N, Corti S, Castellotti B, Mazzini L, Sorarù G, Cereda C, D'Alfonso S, Gellera C, Comi GP, Silani V; SLAGEN Consortium. **(2011)** Novel optineurin mutations in patients with familial and sporadic amyotrophic lateral sclerosis. *J Neurol Neurosurg Psychiatry.* 82(11):1239-43.

Deng HX, Chen W, Hong ST, Boycott KM, Gorrie GH, Siddique N, Yang Y, Fecto F, Shi Y, Zhai H, Jiang H, Hirano M, Rampersaud E, Jansen GH, Donkervoort S, Bigio EH, Brooks BR, Ajroud K, Sufit RL, Haines JL, Mugnaini E, Pericak-Vance MA, Siddique T. **(2011)** Mutations in UBQLN2 cause dominant X-linked juvenile and adult-onset ALS and ALS/dementia. Nature. 477(7363):211-5.

Dice JF. (2007) Chaperone-mediated autophagy. *Autophagy.* 3(4):295-9.

Diekstra FP, van Vught PW, van Rheenen W, Koppers M, Pasterkamp RJ, van Es MA, Schelhaas HJ, de Visser M, Robberecht W, Van Damme P, Andersen PM, van den Berg LH, Veldink JH. **(2012)** UNC13A is a modifier of survival in amyotrophic lateral sclerosis. *Neurobiol Aging.* 33(3):630.e3-8.

Dobson-Stone C, Hallupp M, Bartley L, Shepherd CE, Halliday GM, Schofield PR, Hodges JR, Kwok JB. **(2012)** C9ORF72 repeat expansion in clinical and neuropathologic frontotemporal dementia cohorts. *Neurology.* 79(10):995-1001.

Donnelly CJ, Zhang PW, Pham JT, Haeusler AR, Mistry NA, Vidensky S, Daley EL, Poth EM, Hoover B, Fines DM, Maragakis N, Tienari PJ, Petrucelli L, Traynor BJ, Wang J, Rigo F, Bennett CF, Blackshaw S, Sattler R, Rothstein JD. **(2013)** RNA toxicity from the ALS/FTD C9ORF72 expansion is mitigated by antisense intervention. *Neuron.* 80(2):415-28.

Dooley HC, Razi M, Polson HE, Girardin SE, Wilson MI, Tooze SA. **(2014)** WIPI2 links LC3 conjugation with PI3P, autophagosome formation, and pathogen clearance by recruiting Atg12-5-16L1. *Mol Cell.* 55(2):238-52.

-E-

Elden AC, Kim HJ, Hart MP, Chen-Plotkin AS, Johnson BS, Fang X, Armakola M, Geser F, Greene R, Lu MM, Padmanabhan A, Clay-Falcone D, McCluskey L, Elman L, Juhr D, Gruber PJ, Rüb U, Auburger G, Trojanowski JQ, Lee VM, Van Deerlin VM, Bonini NM, Gitler AD. **(2010)** Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS. *Nature.* 466(7310):1069-75.

-F-

Farg MA, Sundaramoorthy V, Sultana JM, Yang S, Atkinson RA, Levina V, Halloran MA, Gleeson PA, Blair IP, Soo KY, King AE, Atkin JD. **(2014)** C9ORF72, implicated in amytrophic lateral sclerosis and frontotemporal dementia, regulates endosomal trafficking. *Hum Mol Genet.* 23(13):3579-95.

Feng Y, Zhang F, Lokey LK, Chastain JL, Lakkis L, Eberhart D, Warren ST. **(1995)** Translational suppression by trinucleotide repeat expansion at FMR1. *Science.* 268(5211):731-4.

Filimonenko M, Stuffers S, Raiborg C, Yamamoto A, Malerød L, Fisher EM, Isaacs A, Brech A, Stenmark H, Simonsen A. **(2007)** Functional multivesicular bodies are required for autophagic clearance of protein aggregates associated with neurodegenerative disease. *J Cell Biol.* 179(3):485-500.

Filimonenko M, Isakson P, Finley KD, Anderson M, Jeong H, Melia TJ, Bartlett BJ, Myers KM, Birkeland HC, Lamark T, Krainc D, Brech A, Stenmark H, Simonsen A, Yamamoto A. **(2010)** The selective macroautophagic degradation of aggregated proteins requires the PI3P-binding protein Alfy. *Mol Cell.* 38(2):265-79.

Fischbeck KH, Souders D, La Spada A. **(1991)** A candidate gene for X-linked spinal muscular atrophy. *Adv Neurol.* 56:209-13.

Frake RA, Ricketts T, Menzies FM, Rubinsztein DC. **(2015)** Autophagy and neurodegeneration. *J Clin Invest.* 125(1):65-74.

Fratta P, Poulter M, Lashley T, Rohrer JD, Polke JM, Beck J, Ryan N, Hensman D, Mizielinska S, Waite AJ, Lai MC, Gendron TF, Petrucelli L, Fisher EM, Revesz T, Warren JD, Collinge J, Isaacs AM, Mead S. **(2013)** Homozygosity for the C9orf72 GGGGCC repeat expansion in frontotemporal dementia. *Acta Neuropathol.* 126(3):401-9

Freischmidt A, Wieland T, Richter B, Ruf W, Schaeffer V, Müller K, Marroquin N, Nordin F, Hübers A, Weydt P, Pinto S, Press R, Millecamps S, Molko N, Bernard E, Desnuelle C, Soriani MH, Dorst J, Graf E, Nordström U, Feiler MS, Putz S, Boeckers TM, Meyer T, Winkler AS, Winkelman J, de Carvalho M, Thal DR, Otto M, Brännström T, Volk AE, Kursula

P, Danzer KM, Lichtner P, Dikic I, Meitinger T, Ludolph AC, Strom TM, Andersen PM, Weishaupt JH. **(2015)** Haploinsufficiency of TBK1 causes familial ALS and fronto-temporal dementia. *Nat Neurosci*.18(5):631-6.

Freyermuth F, Rau F, Kokunai Y, Linke T, Sellier C, Nakamori M, Kino Y, Arandel L, Jollet A, Thibault C, Philipps M, Vicaire S, Jost B, Udd B, Day JW, Duboc D, Wahbi K, Matsumura T, Fujimura H, Mochizuki H, Deryckere F, Kimura T, Nukina N, Ishiura S, Lacroix V, Campan-Fournier A, Navratil V, Chautard E, Auboeuf D, Horie M, Imoto K, Lee KY, Swanson MS, Lopez de Munain A, Inada S, Itoh H, Nakazawa K, Ashihara T, Wang E, Zimmer T, Furling D, Takahashi MP, Charlet-Berguerand N. **(2016)** Splicing misregulation of SCN5A contributes to cardiac-conduction delay and heart arrhythmia in myotonic dystrophy. *Nat Commun.* 7:11067. doi: 10.1038.

Fu YH, Kuhl DP, Pizzuti A, Pieretti M, Sutcliffe JS, Richards S, Verkerk AJ, Holden JJ, Fenwick RG Jr, Warren ST, et al. (**1991)** Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. *Cell.* 67(6):1047-58.

Fugier C, Klein AF, Hammer C, Vassilopoulos S, Ivarsson Y, Toussaint A, Tosch V, Vignaud A, Ferry A, Messaddeq N, Kokunai Y, Tsuburaya R, de la Grange P, Dembele D, Francois V, Precigout G, Boulade-Ladame C, Hummel MC, Lopez de Munain A, Sergeant N, Laquerrière A, Thibault C, Deryckere F, Auboeuf D, Garcia L, Zimmermann P, Udd B, Schoser B, Takahashi MP, Nishino I, Bassez G, Laporte J, Furling D, Charlet-Berguerand N. **(2011)** Misregulated alternative splicing of BIN1 is associated with T tubule alterations and muscle weakness in myotonic dystrophy. *Nat Med.* 17(6):720-5.

Furukawa K, Kondo T. **(1998)** Identification of the lamina-associated-polypeptide-2-binding domain of B-type lamin. *Eur J Biochem.* 251(3):729-33.

Furukawa K. (1999) LAP2 binding protein 1 (L2BP1/BAF) is a candidate mediator of LAP2-chromatin interaction. *J Cell Sci.* 112 (Pt 15):2485-92.

~G~

Galimberti D, Fenoglio C, Serpente M, Villa C, Bonsi R, Arighi A, Fumagalli GG, Del Bo R, Bruni AC, Anfossi M, Clodomiro A, Cupidi C, Nacmias B, Sorbi S, Piaceri I, Bagnoli S, Bessi V, Marcone A, Cerami C, Cappa SF, Filippi M, Agosta F, Magnani G, Comi G, Franceschi M, Rainero I, Giordana MT, Rubino E, Ferrero P, Rogaeva E, Xi Z, Confaloni A, Piscopo P, Bruno G, Talarico G, Cagnin A, Clerici F, Dell'Osso B, Comi GP, Altamura AC, Mariani C, Scarpini E. **(2013)** Autosomal dominant frontotemporal lobar degeneration due to the C9ORF72 hexanucleotide repeat expansion: late-onset psychotic clinical presentation. *Biol Psychiatry.* 74(5):384-91.

Gami P, Murray C, Schottlaender L, Bettencourt C, De Pablo Fernandez E, Mudanohwo E, Mizielinska S, Polke JM, Holton JL, Isaacs AM, Houlden H, Revesz T, Lashley T. **(2015)** A 30-unit hexanucleotide repeat expansion in C9orf72 induces pathological lesions with dipeptide-repeat proteins and RNA foci, but not TDP-43 inclusions and clinical disease. *Acta Neuropathol.* 130(4):599-601.

Ganley IG, Lam du H, Wang J, Ding X, Chen S, Jiang X. **(2009)** ULK1.ATG13.FIP200 complex mediates mTOR signaling and is essential for autophagy. *J Biol Chem.* 284(18):12297-305.

Gellera C, Tiloca C, Del Bo R, Corrado L, Pensato V, Agostini J, Cereda C, Ratti A, Castellotti B, Corti S, Bagarotti A, Cagnin A, Milani P, Gabelli C, Riboldi G, Mazzini L, Sorarù G, D'Alfonso S, Taroni F, Comi GP, Ticozzi N, Silani V; SLAGEN Consortium. **(2013)** Ubiquilin 2 mutations in Italian patients with amyotrophic lateral sclerosis and frontotemporal dementia. *J Neurol Neurosurg Psychiatry* 84(2):183-7.

Gendron TF, Bieniek KF, Zhang YJ, Jansen-West K, Ash PE, Caulfield T, Daughrity L, Dunmore JH, Castanedes-Casey M, Chew J, Cosio DM, van Blitterswijk M, Lee WC, Rademakers R, Boylan KB, Dickson DW, Petrucelli L. **(2013)** Antisense transcripts of the expanded C9ORF72 hexanucleotide repeat form nuclear RNA foci and undergo repeat-associated non-ATG translation in c9FTD/ALS. *Acta Neuropathol.* 126(6):829-44.

Giannandrea M, Bianchi V, Mignogna ML, Sirri A, Carrabino S, D'Elia E, Vecellio M, Russo S, Cogliati F, Larizza L, Ropers HH, Tzschach A, Kalscheuer V, Oehl-Jaschkowitz B, Skinner C, Schwartz CE, Gecz J, Van Esch H, Raynaud M, Chelly J, de Brouwer AP, Toniolo D, D'Adamo P. **(2010)** Mutations in the small GTPase gene RAB39B are responsible for X-linked mental retardation associated with autism, epilepsy, and macrocephaly. *Am J Hum Genet.* 86(2):185-95.

Gijselinck I, Van Langenhove T, van der Zee J, Sleegers K, Philtjens S, Kleinberger G, Janssens J, Bettens K, Van Cauwenberghe C, Pereson S, Engelborghs S, Sieben A, De Jonghe P, Vandenberghe R, Santens P, De Bleecker J, Maes G, Bäumer V, Dillen L, Joris G, Cuijt I, Corsmit E, Elinck E, Van Dongen J, Vermeulen S, Van den Broeck M, Vaerenberg C, Mattheijssens M, Peeters K, Robberecht W, Cras P, Martin JJ, De Deyn PP, Cruts M, Van Broeckhoven C. **(2012)** A C9orf72 promoter repeat expansion in a Flanders-Belgian cohort with disorders of the frontotemporal lobar degeneration-amyotrophic lateral sclerosis spectrum: a gene identification study. *Lancet Neurol.* 11(1):54-65.

Gijselinck I, Van Mossevelde S, van der Zee J, Sieben A, Engelborghs S, De Bleecker J, Ivanoiu A, Deryck O, Edbauer D, Zhang M, Heeman B, Bäumer V, Van den Broeck M, Mattheijssens M, Peeters K, Rogaeva E, De Jonghe P, Cras P, Martin JJ, de Deyn PP, Cruts M, Van Broeckhoven C. **(2016)** The C9orf72 repeat size correlates with onset age of disease, DNA methylation and transcriptional downregulation of the promoter. *Mol Psychiatry.* 21(8):1112-24.

Gitcho MA, Baloh RH, Chakraverty S, Mayo K, Norton JB, Levitch D, Hatanpaa KJ, White CL 3rd, Bigio EH, Caselli R, Baker M, Al-Lozi MT, Morris JC, Pestronk A, Rademakers R, Goate AM, Cairns NJ. **(2008)** TDP-43 A315T mutation in familial motor neuron disease. *Ann Neurol.* 63(4):535-8.

Gleason CE, Ordureau A, Gourlay R, Arthur JS, Cohen P. **(2011)** Polyubiquitin binding to optineurin is required for optimal activation of TANK-binding kinase 1 and production of interferon β. *J Biol Chem.* 286(41):35663-74.

Gokden M, Al-Hinti JT, Harik SI. **(2009)** Peripheral nervous system pathology in fragile X tremor/ataxia syndrome (FXTAS*). Neuropathology.* 29(3):280-4.

Goldstein LH, Abrahams S. **(2013)** Changes in cognition and behaviour in amyotrophic lateral sclerosis: nature of impairment and implications for assessment. *Lancet Neurol.* 12(4):368-80.

Gomez-Deza J, Lee YB, Troakes C, Nolan M, Al-Sarraj S, Gallo JM, Shaw CE. **(2015)** Dipeptide repeat protein inclusions are rare in the spinal cord and almost absent from motor neurons in C9ORF72 mutant amyotrophic lateral sclerosis and are unlikely to cause their degeneration. *Acta Neuropathol commun.* 3:38.

Gordon PH, Salachas F, Bruneteau G, Pradat PF, Lacomblez L, Gonzalez-Bermejo J, Morelot-Panzini C, Similowski T, Elbaz A, Meininger V. **(2012)** Improving survival in a large French ALS center cohort. *J Neurol.* 259(9):1788-92.

Gorno-Tempini ML, Dronkers NF, Rankin KP, Ogar JM, Phengrasamy L, Rosen HJ, Johnson JK, Weiner MW, Miller BL. **(2004)** Cognition and anatomy in three variants of primary progressive aphasia. *Ann Neurol.* 55(3):335-46.

Grange J, Boyer V, Fabian-Fine R, Fredj NB, Sadoul R, Goldberg Y. **(2004)** Somatodendritic localization and mRNA association of the splicing regulatory protein SAM68 in the hippocampus and cortex. *J Neurosci Res.* 75(5):654-66.

Greco CM, Hagerman RJ, Tassone F, Chudley AE, Del Bigio MR, Jacquemont S, Leehey M, Hagerman PJ. **(2002)** Neuronal intranuclear inclusions in a new cerebellar tremor/ataxia syndrome among fragile X carriers. *Brain.* 125(Pt 8):1760-71.

Greco CM, Berman RF, Martin RM, Tassone F, Schwartz PH, Chang A, Trapp BD, Iwahashi C, Brunberg J, Grigsby J, Hessl D, Becker EJ, Papazian J, Leehey MA, Hagerman RJ, Hagerman PJ. **(2006)** Neuropathology of fragile X-associated tremor/ataxia syndrome (FXTAS). Brain. 129(Pt 1):243-55.

-H-

Haeusler AR, Donnelly CJ, Periz G, Simko EA, Shaw PG, Kim MS, Maragakis NJ, Troncoso JC, Pandey A, Sattler R, Rothstein JD, Wang J. **(2014)** C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature.* 507(7491):195-200.

Haeusler AR, Donnelly CJ, Rothstein JD. **(2016)** The expanding biology of the C9orf72 nucleotide repeat expansion in neurodegenerative disease. *Nat Rev Neurosci.* 17(6):383-95.

Hanada T, Noda NN, Satomi Y, Ichimura Y, Fujioka Y, Takao T, Inagaki F, Ohsumi Y. **(2007)** The Atg12-Atg5 conjugate has a novel E3-like activity for protein lipidation in autophagy. *J Biol Chem.* 282(52):37298-302.

Hagerman RJ, Hagerman PJ. **(2001)** Fragile X syndrome: a model of gene-brain-behaviour relationships. *Rev Neurol.* 33 Suppl 1:S51-7.

Hagerman RJ, Leehey M, Heinrichs W, Tassone F, Wilson R, Hills J, Grigsby J, Gage B, Hagerman PJ. **(2001)** Intention tremor, parkinsonism, and generalized brain atrophy in male carriers of fragile X. *Neurology*. 57(1):127-30.

Hagerman PJ, Hagerman RJ. **(2004)** The fragile-X premutation: a maturing perspective. *Am J Hum Genet. 2004* 805-16.

Hagerman RJ, Leavitt BR, Farzin F, Jacquemont S, Greco CM, Brunberg JA, Tassone F, Hessl D, Harris SW, Zhang L, Jardini T, Gane LW, Ferranti J, Ruiz L, Leehey MA, Grigsby J, Hagerman PJ. **(2004)** Fragile-X-associated tremor/ataxia syndrome (FXTAS) in females with the FMR1 premutation. *Am J Hum Genet.* 74(5):1051-6.

Hagerman PJ. (2008) The fragile X prevalence paradox. *J Med Genet*. 45(8):498-9.

Hara T, Nakamura K, Matsui M, Yamamoto A, Nakahara Y, Suzuki-Migishima R, Yokoyama M, Mishima K, Saito I, Okano H, Mizushima N. **(2006)** Suppression of basal autophagy in neural cells causes neurodegenerative disease in mice. *Nature.* 441(7095):885-9.

Hart MP, Brettschneider J, Lee VM, Trojanowski JQ, Gitler AD. **(2012)** Distinct TDP-43 pathology in ALS patients with ataxin 2 intermediate-length polyQ expansions. *Acta Neuropathol.* 124(2):221-30.

Heo JM, Ordureau A, Paulo JA, Rinehart J, Harper JW. **(2015)** The PINK1-PARKIN Mitochondrial Ubiquitylation Pathway Drives a Program of OPTN/NDP52 Recruitment and TBK1 Activation to Promote Mitophagy. *Mol Cell.* 60(1):7-20.

Hessl D, Tassone F, Loesch DZ, Berry-Kravis E, Leehey MA, Gane LW, Barbato I, Rice C, Gould E, Hall DA, Grigsby J, Wegelin JA, Harris S, Lewin F, Weinberg D, Hagerman PJ, Hagerman RJ. **(2005)** Abnormal elevation of FMR1 mRNA is associated with psychological symptoms in individuals with the fragile X premutation. *Am J Med Genet B Neuropsychiatr Genet.* 139B(1):115-21.

Hinds HL, Ashley CT, Sutcliffe JS, Nelson DL, Warren ST, Housman DE, Schalling M. **(1993)** Tissue specific expression of FMR-1 provides evidence for a functional role in fragile X syndrome. *Nat Genet.* 3(1):36-43.

Hirano M, Nakamura Y, Saigoh K, Sakamoto H, Ueno S, Isono C, Miyamoto K, Akamatsu M, Mitsui Y, Kusunoki S. **(2013)** Mutations in the gene encoding p62 in Japanese patients with amyotrophic lateral sclerosis. *Neurology.* 80(5):458-63.

Hodges JR, Davies R, Xuereb J, Kril J, Halliday G. **(2003)** Survival in frontotemporal dementia. *Neurology.* 61(3):349-54.

Hukema RK, Buijsen RA, Schonewille M, Raske C, Severijnen LA, Nieuwenhuizen-Bakker I, Verhagen RF, van Dessel L, Maas A, Charlet-Berguerand N, De Zeeuw CI, Hagerman PJ, Berman RF, Willemsen R. **(2015)** Reversibility of neuropathology and motor deficits in an inducible mouse model for FXTAS. *Hum Mol Genet.* 24(17):4948-57.

Hunsaker MR, Greco CM, Spath MA, Smits AP, Navarro CS, Tassone F, Kros JM, Severijnen LA, Berry-Kravis EM, Berman RF, Hagerman PJ, Willemsen R, Hagerman RJ, Hukema RK. **(2011)** Widespread non-central nervous system organ pathology in fragile X premutation carriers with fragile X-associated tremor/ataxia syndrome and CGG knock-in mice*. Acta Neuropathol.* 122(4):467-79.

Hsu PP, Kang SA, Rameseder J, Zhang Y, Ottina KA, Lim D, Peterson TR, Choi Y, Gray NS, Yaffe MB, Marto JA, Sabatini DM. **(2011)** The mTOR-regulated phosphoproteome reveals a mechanism of mTORC1-mediated inhibition of growth factor signaling. *Science.* 332(6035):1317-22.

-I-

Imbert G, Saudou F, Yvert G, Devys D, Trottier Y, Garnier JM, Weber C, Mandel JL, Cancel G, Abbas N, Dürr A, Didierjean O, Stevanin G, Agid Y, Brice A. **(1996)** Cloning of the gene for spinocerebellar ataxia 2 reveals a locus with high sensitivity to expanded CAG/glutamine repeats. *Nat Genet.* 14(3):285-91.

Ingolia NT, Lareau LF, Weissman JS**. (2011)** Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell.* 147(4):789-802.

Iwahashi CK, Yasui DH, An HJ, Greco CM, Tassone F, Nannen K, Babineau B, Lebrilla CB, Hagerman RJ, Hagerman PJ. **(2006)** Protein composition of the intranuclear inclusions of FXTAS. *Brain*. 129(Pt 1):256-71. Epub 2005 Oct 24.

-J-

Jacquemont S, Hagerman RJ, Leehey M, Grigsby J, Zhang L, Brunberg JA, Greco C, Des Portes V, Jardini T, Levine R, Berry-Kravis E, Brown WT, Schaeffer S, Kissel J, Tassone F, Hagerman PJ. **(2003)** Fragile X premutation tremor/ataxia syndrome: molecular, clinical, and neuroimaging correlates. Am J Hum Genet. 72(4):869-78.

Jacquemont S (a), Hagerman RJ, Leehey MA, Hall DA, Levine RA, Brunberg JA, Zhang L, Jardini T, Gane LW, Harris SW, Herman K, Grigsby J, Greco CM, Berry-Kravis E, Tassone F, Hagerman PJ. **(2004)** Penetrance of the fragile X-associated tremor/ataxia syndrome in a premutation carrier population. *JAMA.* 291(4):460-9.

Jacquemont S (b), Farzin F, Hall D, Leehey M, Tassone F, Gane L, Zhang L, Grigsby J, Jardini T, Lewin F, Berry-Kravis E, Hagerman PJ, Hagerman RJ. **(2004)** Aging in individuals with the FMR1 mutation. *Am J Ment Retard.* 109(2):154-64.

Jiang J, Zhu Q, Gendron TF, Saberi S, McAlonis-Downes M, Seelman A, Stauffer JE, Jafar-Nejad P, Drenner K, Schulte D, Chun S, Sun S, Ling SC, Myers B, Engelhardt J, Katz M, Baughn M, Platoshyn O, Marsala M, Watt A, Heyser CJ, Ard MC, De Muynck L, Daughrity LM, Swing DA, Tessarollo L, Jung CJ, Delpoux A, Utzschneider DT, Hedrick SM, de Jong PJ, Edbauer D, Van Damme P, Petrucelli L, Shaw CE, Bennett CF, Da Cruz S, Ravits J, Rigo F, Cleveland DW, Lagier-Tourenne C. **(2016)** Gain of Toxicity from ALS/FTD-Linked Repeat Expansions in C9ORF72 Is Alleviated by Antisense Oligonucleotides Targeting GGGGCC-Containing RNAs. *Neuron.* 90(3):535-50.

Jin P, Duan R, Qurashi A, Qin Y, Tian D, Rosser TC, Liu H, Feng Y, Warren ST. **(2007)** Pur alpha binds to rCGG repeats and modulates repeat-mediated neurodegeneration in a Drosophila model of fragile X tremor/ataxia syndrome. *Neuron*. 55(4):556-64.

Johnson JO, Mandrioli J, Benatar M, Abramzon Y, Van Deerlin VM, Trojanowski JQ, Gibbs JR, Brunetti M, Gronka S, Wuu J, Ding J, McCluskey L, Martinez-Lage M, Falcone D, Hernandez DG, Arepalli S, Chong S, Schymick JC, Rothstein J, Landi F, Wang YD, Calvo A, Mora G, Sabatelli M, Monsurrò MR, Battistini S, Salvi F, Spataro R, Sola P, Borghero G; ITALSGEN Consortium, Galassi G, Scholz SW, Taylor JP, Restagno G, Chiò A, Traynor BJ. **(2010)** Exome sequencing reveals VCP mutations as a cause of familial ALS. *Neuron.* 68(5):857-64.

Johnson CW, Melia TJ, Yamamoto A. **(2012)** Modulating macroautophagy: a neuronal perspective. *Future Med Chem.* 4(13):1715-31.

Johnson JO (a), Glynn SM, Gibbs JR, Nalls MA, Sabatelli M, Restagno G, Drory VE, Chiò A, Rogaeva E, Traynor BJ. **(2014)** Mutations in the CHCHD10 gene are a common cause of familial amyotrophic lateral sclerosis. *Brain*. 2014;137(pt 12):e311.

Johnson JO (b), Pioro EP, Boehringer A, Chia R, Feit H, Renton AE, Pliner HA, Abramzon Y, Marangi G, Winborn BJ, Gibbs JR, Nalls MA, Morgan S, Shoai M, Hardy J, Pittman A, Orrell RW, Malaspina A, Sidle KC, Fratta P, Harms MB, Baloh RH, Pestronk A, Weihl CC, Rogaeva E, Zinman L, Drory VE, Borghero G, Mora G, Calvo A, Rothstein JD; ITALSGEN Consortium, Drepper C, Sendtner M, Singleton AB, Taylor JP, Cookson MR, Restagno G, Sabatelli M, Bowser R, Chiò A, Traynor BJ. **(2014)** Mutations in the Matrin 3 gene cause familial amyotrophic lateral sclerosis. *Nat Neurosci.* 17(5):664-6.

Jongens TA, Ackerman LD, Swedlow JR, Jan LY, Jan YN. **(1994)** Germ cell-less encodes a cell type-specific nuclear pore-associated protein and functions early in the germ-cell specification pathway of Drosophila. *Genes Dev.* 8(18):2123-36.

Josephs KA, Hodges JR, Snowden JS, Mackenzie IR, Neumann M, Mann DM, Dickson DW. **(2011)** Neuropathological background of phenotypical variability in frontotemporal dementia. *Acta Neuropathol.* 122(2):137-53.

Jovičić A, Mertens J, Boeynaems S, Bogaert E, Chai N, Yamada SB, Paul JW 3rd, Sun S, Herdy JR, Bieri G, Kramer NJ, Gage FH, Van Den Bosch L, Robberecht W, Gitler AD. **(2015)** Modifiers of C9orf72 dipeptide repeat toxicity connect nucleocytoplasmic transport defects to FTD/ALS. *Nat Neurosci.* 18(9):1226-9.

Ju JS, Miller SE, Jackson E, Cadwell K, Piwnica-Worms D, Weihl CC. **(2009)** Quantitation of selective autophagic protein aggregate degradation in vitro and in vivo using luciferase reporters. *Autophagy.* 5(4):511-9. Epub 2009 May 6.

-K-

Kabashi E, Valdmanis PN, Dion P, Spiegelman D, McConkey BJ, Vande Velde C, Bouchard JP, Lacomblez L, Pochigaeva K, Salachas F, Pradat PF, Camu W, Meininger V, Dupre N, Rouleau GA. **(2008)** TARDBP mutations in individuals with sporadic and familial amyotrophic lateral sclerosis. *Nat Genet.* 40(5):572-4.

Kachaner D, Génin P, Laplantine E, Weil R. **(2012)** Toward an integrative view of Optineurin functions. *Cell Cycle.* 11(15):2808-18.

Kaiyala KJ, Schwartz MW. **(2011)** Toward a more complete (and less controversial) understanding of energy expenditure and its role in obesity pathogenesis. *Diabetes.* 60(1):17-23.

Kamm C, Healy DG, Quinn NP, Wüllner U, Moller JC, Schols L, Geser F, Burk K, Børglum AD, Pellecchia MT, Tolosa E, del Sorbo F, Nilsson C, Bandmann O, Sharma M, Mayer P, Gasteiger M, Haworth A, Ozawa T, Lees AJ, Short J, Giunti P, Holinski-Feder E, Illig T, Wichmann HE, Wenning GK, Wood NW, Gasser T; European Multiple System Atrophy Study Group. **(2005)** The fragile X tremor ataxia syndrome in the differential diagnosis of multiple system atrophy: data from the EMSA Study Group. *Brain.* 128(Pt 8):1855-60.

Kearse MG, Green KM, Krans A, Rodriguez CM, Linsalata AE, Goldstrohm AC, Todd PK. **(2016)** CGG Repeat-Associated Non-AUG Translation Utilizes a Cap-Dependent Scanning Mechanism of Initiation to Produce Toxic Proteins. *Mol Cell.* 62(2):314-22.

Kenneson A, Zhang F, Hagedorn CH, Warren ST. **(2001)** Reduced FMRP and increased FMR1 transcription is proportionally associated with CGG repeat number in intermediate-length and premutation carriers. *Hum Mol Genet.* 10(14):1449-54.

Kim EJ, Kwon JC, Park KH, Park KW, Lee JH, Choi SH, Jeong JH, Kim BC, Yoon SJ, Yoon YC, Kim S, Park KC, Choi BO, Na DL, Ki CS, Kim SH. **(2014)** Clinical and genetic analysis of MAPT, GRN, and C9orf72 genes in Korean patients with frontotemporal dementia. *Neurobiol Aging.* 35(5):1213.e13-7.

Kirkin V, McEwan DG, Novak I, Dikic I. **(2009)** A role for ubiquitin in selective autophagy. *Mol Cell.* 34(3):259-69.

Klionsky DJ, Emr SD. **(2000)** Autophagy as a regulated pathway of cellular degradation. *Science.* 2000 290(5497):1717-21.

Klionsky DJ. (2007) Autophagy: from phenomenology to molecular understanding in less than a decade. *Nat Rev Mol Cell Biol.* 8(11):931-7.

Komatsu M, Waguri S, Chiba T, Murata S, Iwata J, Tanida I, Ueno T, Koike M, Uchiyama Y, Kominami E, Tanaka K. **(2006)** Loss of autophagy in the central nervous system causes neurodegeneration in mice. *Nature.* 441(7095):880-4. Epub 2006 Apr 19.

Koppers M, Blokhuis AM, Westeneng HJ, Terpstra ML, Zundel CA, Vieira de Sá R, Schellevis RD, Waite AJ, Blake DJ, Veldink JH, van den Berg LH, Pasterkamp RJ. **(2015)** C9orf72 ablation in mice does not cause motor neuron degeneration or motor deficits. *Ann Neurol.* 2015 Sep;78(3):426-38.

Kostić VS, Dobričić V, Stanković I, Ralić V, Stefanova E. **(2014)** C9orf72 expansion as a possible genetic cause of Huntington disease phenocopy syndrome. *J Neurol.* 261(10):1917-21.

Kovacs GG, van der Zee J, Hort J, Kristoferitsch W, Leitha T, Höftberger R, Ströbel T, Van Broeckhoven C, Matej R.(2016) Clinicopathological description of two cases with SQSTM1 gene mutation associated with frontotemporal dementia. Neuropathology. 36(1):27-38.

Kowalczyk MS, Hughes JR, Babbs C, Sanchez-Pulido L, Szumska D, Sharpe JA, Sloane-Stanley JA, Morriss-Kay GM, Smoot LB, Roberts AE, Watkins H, Bhattacharya S, Gibbons RJ, Ponting CP, Wood WG, Higgs DR. (**2012)** Nprl3 is required for normal development of the cardiovascular system. *Mamm Genome.* 23(7-8):404-15.

Kozak M. (1988) Leader length and secondary structure modulate mRNA function under conditions of stress. *Mol Cell Biol.* 8(7):2737-44.

Kozak M. (1990) Downstream secondary structure facilitates recognition of initiator codons by eukaryotic ribosomes. *Proc Natl Acad Sci U S A.* 87(21):8301-5.

Kremer B, Goldberg P, Andrew SE, Theilmann J, Telenius H, Zeisler J, Squitieri F, Lin B, Bassett A, Almqvist E, et al. **(1994)** A worldwide study of the Huntington's disease mutation. The sensitivity and specificity of measuring CAG repeats. *N Engl J Med.* 330(20):1401-6.

Kwon I, Xiang S, Kato M, Wu L, Theodoropoulos P, Wang T, Kim J, Yun J, Xie Y, McKnight SL. **(2014)** Poly-dipeptides encoded by the C9orf72 repeats bind nucleoli, impede RNA biogenesis, and kill cells. *Science.* 345(6201):1139-45.

-L-

Lagier-Tourenne C, Baughn M, Rigo F, Sun S, Liu P, Li HR, Jiang J, Watt AT, Chun S, Katz M, Qiu J, Sun Y, Ling SC, Zhu Q, Polymenidou M, Drenner K, Artates JW, McAlonis-Downes M, Markmiller S, Hutt KR, Pizzo DP, Cady J, Harms MB, Baloh RH, Vandenberg SR, Yeo GW, Fu XD, Bennett CF, Cleveland DW, Ravits J. **(2013)** Targeted degradation of sense and antisense C9orf72 RNA foci as therapy for ALS and frontotemporal degeneration. *Proc Natl Acad Sci U S A.* 110(47):E4530-9.

Lattante S, Millecamps S, Stevanin G, Rivaud-Péchoux S, Moigneu C, Camuzat A, Da Barroca S, Mundwiller E, Couarch P, Salachas F, Hannequin D, Meininger V, Pasquier F, Seilhean D, Couratier P, Danel-Brunaud V, Bonnet AM, Tranchant C, LeGuern E, Brice A, Le Ber I, Kabashi E; French Research Network on FTD and FTD-ALS. **(2014)** Contribution of ATXN2 intermediary polyQ expansions in a spectrum of neurodegenerative disorders. *Neurology.* 83(11):990-5.

Lazarou M, Sliter DA, Kane LA, Sarraf SA, Wang C, Burman JL, Sideris DP, Fogel AI, Youle RJ. **(2015)** The ubiquitin kinase PINK1 recruits autophagy receptors to induce mitophagy. *Nature.* 524(7565):309-14.

Le Ber I, Camuzat A, Guerreiro R, Bouya-Ahmed K, Bras J, Nicolas G, Gabelle A, Didic M, De Septenville A, Millecamps S, Lenglet T, Latouche M, Kabashi E, Campion D, Hannequin D, Hardy J, Brice A; French Clinical and Genetic Research Network on FTD/FTD-ALS. **(2013)** *JAMA Neurol.* 70(11):1403-10.

Lee KK, Haraguchi T, Lee RS, Koujin T, Hiraoka Y, Wilson KL. **(2001)** Distinct functional domains in emerin bind lamin A and DNA-bridging protein BAF. *J Cell Sci.* 114(Pt 24):4567-73.

Lee Y, Ahn C, Han J, Choi H, Kim J, Yim J, Lee J, Provost P, Rådmark O, Kim S, Kim VN. **(2003)** The nuclear RNase III Drosha initiates microRNA processing. *Nature.* 425(6956):415-9.

Lee JE, Cooper TA. **(2009)** Pathogenic mechanisms of myotonic dystrophy. *Biochem Soc Trans.* 37(Pt 6):1281-6. doi: 10.1042/BST0371281.

Lee T, Li YR, Ingre C, Weber M, Grehl T, Gredal O, de Carvalho M, Meyer T, Tysnes OB, Auburger G, Gispert S, Bonini NM, Andersen PM, Gitler AD. **(2011)** Ataxin-2 intermediate-length polyglutamine expansions in European ALS patients. Hum Mol Genet. 20(9):1697-700.

Lee S, Liu B, Lee S, Huang SX, Shen B, Qian SB. **(2012)** Global mapping of translation initiation sites in mammalian cells at single-nucleotide resolution. *Proc Natl Acad Sci U S A.* 109(37):E2424-32.

Lee SE (a), Tartaglia MC, Yener G, Genç S, Seeley WW, Sanchez-Juan P, Moreno F, Mendez MF, Klein E, Rademakers R, López de Munain A, Combarros O, Kramer JH, Kenet RO, Boxer AL, Geschwind MD, Gorno-Tempini ML, Karydas AM, Rabinovici GD, Coppola G, Geschwind DH, Miller BL. **(2013)** Neurodegenerative disease phenotypes in carriers of MAPT p.A152T, a risk factor for frontotemporal dementia spectrum disorders and Alzheimer disease. *Alzheimer Dis Assoc Disord.* 27(4):302-9.

Lee YB (b), Chen HJ, Peres JN, Gomez-Deza J, Attig J, Stalekar M, Troakes C, Nishimura AL, Scotter EL, Vance C, Adachi Y, Sardone V, Miller JW, Smith BN, Gallo JM, Ule J, Hirth F, Rogelj B, Houart C, Shaw CE. **(2013)** Hexanucleotide repeats in ALS/FTD form length-dependent RNA foci, sequester RNA binding proteins, and are neurotoxic. *Cell Rep.* 5(5):1178-86.

Leehey MA, Berry-Kravis E, Min SJ, Hall DA, Rice CD, Zhang L, Grigsby J, Greco CM, Reynolds A, Lara R, Cogswell J, Jacquemont S, Hessl DR, Tassone F, Hagerman R, Hagerman PJ. **(2007)** Progression of tremor and ataxia in male carriers of the FMR1 premutation. *Mov Disord.* 22(2):203-6.

Leigh PN, Whitwell H, Garofalo O, Buller J, Swash M, Martin JE, Gallo JM, Weller RO, Anderton BH. **(1991)** Ubiquitin-immunoreactive intraneuronal inclusions in amyotrophic lateral sclerosis. Morphology, distribution, and specificity. *Brain.* 114 (Pt 2):775-88.

Leibel RL. (2008) Molecular physiology of weight regulation in mice and humans. *Int J Obes (Lond).* 32 Suppl 7:S98-108.

Levine TP, Daniels RD, Gatta AT, Wong LH, Hayes MJ. **(2013)** The product of C9orf72, a gene strongly implicated in neurodegeneration, is structurally related to DENN Rab-GEFs. *Bioinformatics.* 29(4):499-503.

Li WW, Li J, Bao JK. **(2012)** Microautophagy: lesser-known self-eating. *Cell Mol Life Sci.* 69(7):1125-36.

Lin X, Miller JW, Mankodi A, Kanadia RN, Yuan Y, Moxley RT, Swanson MS, Thornton CA. **(2006)** Failure of MBNL1-dependent postnatal splicing transitions in myotonic dystrophy. *Hum Mol Genet.* 15:2087–2097.

Lindquist SG, Braedgaard H, Svenstrup K, Isaacs AM, Nielsen JE; FReJA Consortium. **(2008)** Frontotemporal dementia linked to chromosome 3 (FTD-3)--current concepts and the detection of a previously unknown branch of the Danish FTD-3 family. *Eur J Neurol.* 15(7):667-70.

Lipton AM, White CL, Bigio EH. **(2004)** Frontotemporal lobar degeneration with motor neuron disease-type inclusions predominates in 76 cases of frontotemporal degeneration. *Acta Neuropathol.* 108(5):379-85.

Liu F, Liu Q, Lu CX, Cui B, Guo XN, Wang RR, Liu MS, Li XG, Cui LY, Zhang X. **(2016)** Identification of a novel loss-of-function C9orf72 splice site mutation in a patient with amyotrophic lateral sclerosis. *Neurobiol Aging.* pii: S0197-4580(16)30164-6.

Loesch DZ, Churchyard A, Brotchie P, Marot M, Tassone F. **(2005)** Evidence for, and a spectrum of, neurological involvement in carriers of the fragile X pre-mutation: FXTAS and beyond.*Clin Genet.* 67(5):412-7.

Logroscino G, Traynor BJ, Hardiman O, Chiò A, Mitchell D, Swingler RJ, Millul A, Benn E, Beghi E; EURALS. **(2010)** Incidence of amyotrophic lateral sclerosis in Europe. *J Neurol Neurosurg Psychiatry.* 81(4):385-90.

Lomen-Hoerth C, Murphy J, Langmore S, Kramer JH, Olney RK, Miller B. **(2003)** Are amyotrophic lateral sclerosis patients cognitively normal? *Neurology.* 60(7):1094-7.

Loureiro JR, Oliveira CL, Silveira I **(2016)** Unstable repeat expansions in neurodegenerative diseases: nucleocytoplasmic transport emerges on the scene. *Neurobiol Aging*. 39:174-83.

Lukong KE, Richard S. **(2008)** Motor coordination defects in mice deficient for the Sam68 RNA-binding protein. *Behav Brain Res.* 189(2):357-63.

~M~

McGuire D, Garrison L, Armon C, Barohn R, Bryan W, Miller R, Parry G, Petajan J, Ross M. **(1996)** Relationship of the Tufts Quantitative Neuromuscular Exam (TQNE) and the Sickness Impact Profile (SIP) in measuring progression of ALS. SSNJV/CNTF ALS Study Group. *Neurology.* 46(5):1442-4.

Mackenzie IR, Arzberger T, Kremmer E, Troost D, Lorenz S, Mori K, Weng SM, Haass C, Kretzschmar HA, Edbauer D, Neumann M. **(2013)** Dipeptide repeat protein pathology in C9ORF72 mutation cases: clinico-pathological correlations. *Acta neuropathologica* 126 ; 859-879.

Mackenzie IR, Frick P, Grässer FA, Gendron TF, Petrucelli L, Cashman NR, Edbauer D, Kremmer E, Prudlo J, Troost D, Neumann M. (2015) Quantitative analysis and clinico-pathological correlations of different dipeptide repeat protein pathologies in C9ORF72 mutation carriers. *Acta Neuropathol.* 130(6):845-61.

Majounie E, Renton AE, Mok K, Dopper EG, Waite A, Rollinson S, Chiò A, Restagno G, Nicolaou N, Simon-Sanchez J, van Swieten JC, Abramzon Y, Johnson JO, Sendtner M, Pamphlett R, Orrell RW, Mead S, Sidle KC, Houlden H, Rohrer JD, Morrison KE, Pall H, Talbot K, Ansorge O; Chromosome 9-ALS/FTD Consortium; French research network on FTLD/FTLD/ALS; ITALSGEN Consortium, Hernandez DG, Arepalli S, Sabatelli M, Mora G, Corbo M, Giannini F, Calvo A, Englund E, Borghero G, Floris GL, Remes AM, Laaksovirta H, McCluskey L, Trojanowski JQ, Van Deerlin VM, Schellenberg GD, Nalls MA, Drory VE, Lu CS, Yeh TH, Ishiura H, Takahashi Y, Tsuji S, Le Ber I, Brice A, Drepper C, Williams N, Kirby J, Shaw P, Hardy J, Tienari PJ, Heutink P, Morris HR, Pickering-Brown S, Traynor BJ. **(2012)** Frequency of the C9orf72 hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study. *Lancet Neurol.* 11(4):323-30.

Mann DM, Rollinson S, Robinson A, Bennion Callister J, Thompson JC, Snowden JS, Gendron T, Petrucelli L, Masuda-Suzukake M, Hasegawa M, Davidson Y, Pickering-Brown S. **(2013)** Dipeptide repeat proteins are present in the p62 positive inclusions in patients with frontotemporal lobar degeneration and motor neurone disease associated with expansions in C9ORF72. *Acta Neuropathol Commun.* 1 : 68.

Mankodi A, Thornton CA. **(2002)** Myotonic syndromes. *Curr Opin Neurol.* 15(5):545-52.

Marambio P, Toro B, Sanhueza C, Troncoso R, Parra V, Verdejo H, García L, Quiroga C, Munafo D, Díaz-Elizondo J, Bravo R, González MJ, Diaz-Araya G, Pedrozo Z, Chiong M, Colombo MI, Lavandero S. **(2010)** Glucose deprivation causes oxidative stress and stimulates aggresome formation and autophagy in cultured cardiac myocytes*. Biochim Biophys Acta.* 1802(6):509-18.

Marangi G, Traynor BJ. **(2015)** Genetic causes of amyotrophic lateral sclerosis: new genetic analysis methodologies entailing new opportunities and challenges. *Brain Res.* 1607:75-93.

Marat AL, Dokainish H, McPherson PS. **(2011)** DENN domain proteins: regulators of Rab GTPases. *J Biol Chem.* 286(16):13791-800.

Maruyama H, Morino H, Ito H, Izumi Y, Kato H, Watanabe Y, Kinoshita Y, Kamada M, Nodera H, Suzuki H, Komure O, Matsuura S, Kobatake K, Morimoto N, Abe K, Suzuki N, Aoki M, Kawata A, Hirai T, Kato T, Ogasawara K, Hirano A, Takumi T, Kusaka H, Hagiwara K, Kaji R, Kawakami H. **(2010)** Mutations of optineurin in amyotrophic lateral sclerosis. *Nature.* 465(7295):223-6.

Maruyama H, Kawakami H. **(2013)** Optineurin and amyotrophic lateral sclerosis. *Geriatr Gerontol Int.* 13(3):528-32.

Matilla-Dueñas A, Ashizawa T, Brice A, Magri S, McFarland KN, Pandolfo M, Pulst SM, Riess O, Rubinsztein DC, Schmidt J, Schmidt T, Scoles DR, Stevanin G, Taroni F, Underwood BR, Sánchez I. **(2014)** Consensus paper: pathological mechanisms underlying neurodegeneration in spinocerebellar ataxias. *Cerebellum.* 13(2):269-302.

Matsumoto G, Shimogori T, Hattori N, Nukina N. **(2015)** TBK1 controls autophagosomal engulfment of polyubiquitinated mitochondria through p62/SQSTM1 phosphorylation. *Hum Mol Genet.* 24(15):4429-42

Millecamps S, De Septenville A, Teyssou E, Daniau M, Camuzat A, Albert M, LeGuern E, Galimberti D; French research network on FTD and FTD-ALS, Brice A, Marie Y, Le Ber I. **(2014)** Genetic analysis of matrin 3 gene in French amyotrophic lateral sclerosis patients and frontotemporal lobar degeneration with amyotrophic lateral sclerosis patients. *Neurobiol Aging.* 35(12):2882.e13-5.

Miller JW, Urbinati CR, Teng-umnuay P, Stenberg MG, Byrne BJ, Thornton CA, Swanson MS. **(2000)** Recruitment of human muscleblind proteins to (CUG)(n) expansions associated with myotonic dystrophy. *EMBO J.* 19:4439–4448.

Mizielinska S, Lashley T, Norona FE, Clayton EL, Ridler CE, Fratta P, Isaacs AM. **(2013)** C9orf72 frontotemporal lobar degeneration is characterised by frequent neuronal sense and antisense RNA foci. *Acta Neuropathol.* 126(6):845-57.

Mizielinska S, Grönke S, Niccoli T, Ridler CE, Clayton EL, Devoy A, Moens T, Norona FE, Woollacott IO, Pietrzyk J, Cleverley K, Nicoll AJ, Pickering-Brown S, Dols J, Cabecinha M, Hendrich O, Fratta P, Fisher EM, Partridge L, Isaacs

AM. **(2014)** C9orf72 repeat expansions cause neurodegeneration in Drosophila through arginine-rich proteins. *Science.* 345(6201):1192-4.

Mizushima N (2007) Autophagy: process and function. *Genes Dev.* 21(22):2861-73.

Mori K (a), Arzberger T., Grasser F.A., Gijselinck I., May S., Rentzsch K., Weng S.M., Schludi M.H., van der Zee J., Cruts M., Van Broeckhoven C., Kremmer E., Kretzschmar H.A., Haass C., Edbauer D. **(2013)** Bidirectional transcripts of the expanded C9orf72 hexanucleotide repeat are translated into aggregating dipeptide repeat proteins. *Acta neuropathologica.* 126 ; 881-893.

Mori M (b), Weng S.M., Arzberger T., May S., Rentzsch K., Kremmer E., Schmid B., Kretzschmar H.A., Cruts M., Van Broeckhoven C., Haass C., Edbauer D., **(2013)** The C9orf72 GGGGCC repeat is translated into aggregating dipeptide repeat proteins in FTLD/ALS. *Science* 339 ; 1335 -1338.

Morton S, Hesson L, Peggie M, Cohen P. **(2008)** Enhanced binding of TBK1 by an optineurin mutant that causes a familial form of primary open angle glaucoma. *FEBS Lett.* 582(6):997-1002.

Murphy J, Henry R, Lomen-Hoerth C. **(2007)** Establishing subtypes of the continuum of frontal lobe impairment in amyotrophic lateral sclerosis. *Arch Neurol.* 64(3):330-4.

-N-

Nakatogawa H, Ichimura Y, Ohsumi Y. **(2007)** Atg8, a ubiquitin-like protein required for autophagosome formation, mediates membrane tethering and hemifusion. *Cell.* 130(1):165-78.

Neary D, Snowden JS, Gustafson L, Passant U, Stuss D, Black S, Freedman M, Kertesz A, Robert PH, Albert M, Boone K, Miller BL, Cummings J, Benson DF. **(1998)** Frontotemporal lobar degeneration: a consensus on clinical diagnostic criteria. *Neurology.* 51(6):1546-54.

Neumann M, Sampathu DM, Kwong LK, Truax AC, Micsenyi MC, Chou TT, Bruce J, Schuck T, Grossman M, Clark CM, McCluskey LF, Miller BL, Masliah E, Mackenzie IR, Feldman H, Feiden W, Kretzschmar HA, Trojanowski JQ, Lee VM. **(2006)** Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science.* 314(5796):130-3.

Nili E, Cojocaru GS, Kalma Y, Ginsberg D, Copeland NG, Gilbert DJ, Jenkins NA, Berger R, Shaklai S, Amariglio N, Brok-Simoni F, Simon AJ, Rechavi G. **(2001)** Nuclear membrane protein LAP2beta mediates transcriptional repression alone and together with its binding partner GCL (germ-cell-less). *J Cell Sci.* 114(Pt 18):3297-307.

Nixon RA, **(2013)** The role of autophagy in neurodegenrative disease. *Nat Med.* 19(8):983-97.

Niwa H, Yamamura K, Miyazaki J. **(1991)** Efficient selection for high-expression transfectants with a novel eukaryotic factor. *Gene.* 108(2):193-9.

Nordquist DT, Kozak CA, Orr HT. **(1988)**. cDNA cloning and characterization of three genes uniquely expressed in cerebellum by Purkinje neurons. *J Neurosci* 8:4780 – 4789.

-O-

Oberdick J, Smeyne RJ, Mann JR, Zackson S, Morgan JI. **(1990)**. A promoter that drives transgene expression in cerebellar Purkinje and retinal bipolar neurons. **Science** 248:223–226.

Oberlé I, Rousseau F, Heitz D, Kretz C, Devys D, Hanauer A, Boué J, Bertheas MF, Mandel JL. **(1991)** Instability of a 550-base pair DNA segment and abnormal methylation in fragile X syndrome. *Science* 252(5009):1097-102.

O'Rourke JG, Bogdanik L, Muhammad AK, Gendron TF, Kim KJ, Austin A, Cady J, Liu EY, Zarrow J, Grant S, Ho R, Bell S, Carmona S, Simpkinson M, Lall D, Wu K, Daughrity L, Dickson DW, Harms MB, Petrucelli L, Lee EB, Lutz CM, Baloh RH. **(2015)** C9orf72 BAC Transgenic Mice Display Typical Pathologic Features of ALS/FTD. *Neuron.* 2015 Dec 2;88(5):892-901.

O'Rourke JG, Bogdanik L, Yáñez A, Lall D, Wolf AJ, Muhammad AK, Ho R, Carmona S, Vit JP, Zarrow J, Kim KJ, Bell S, Harms MB, Miller TM, Dangler CA, Underhill DM, Goodridge HS, Lutz CM, Baloh RH. **(2016)** C9orf72 is required for proper macrophage and microglial function in mice. *Science.* 351(6279):1324-9.

-P-

Parkinson N, Ince PG, Smith MO, Highley R, Skibinski G, Andersen PM, Morrison KE, Pall HS, Hardiman O, Collinge J, Shaw PJ, Fisher EM; MRC Proteomics in ALS Study; FReJA Consortium. **(2006)** ALS phenotypes with mutations in CHMP2B (charged multivesicular body protein 2B). *Neurology.* 67(6):1074-7. E

Paronetto MP, Achsel T, Massiello A, Chalfant CE, Sette C. **(2007)** The RNA-binding protein Sam68 modulates the alternative splicing of Bcl-x. *J Cell Biol.* 176(7):929-39.

Pastula DM, Moore DH, Bedlack RS. **(2010)** Creatine for amyotrophic lateral sclerosis/motor neuron disease. *Cochrane Database Syst Rev* (6):CD005225

Pedrotti S, Bielli P, Paronetto MP, Ciccosanti F, Fimia GM, Stamm S, Manley JL, Sette C. **(2010)** The splicing regulator Sam68 binds to a novel exonic splicing silencer and functions in SMN2 alternative splicing in spinal muscular atrophy. *EMBO J.* 29(7):1235-47.

Peters OM, Cabrera GT, Tran H, Gendron TF, McKeon JE, Metterville J, Weiss A, Wightman N, Salameh J, Kim J, Sun H, Boylan KB, Dickson D, Kennedy Z, Lin Z, Zhang YJ, Daughrity L, Jung C, Gao FB, Sapp PC, Horvitz HR, Bosco DA, Brown SP, de Jong P, Petrucelli L, Mueller C, Brown RH Jr. **(2015)** Human C9ORF72 Hexanucleotide Expansion Reproduces RNA Foci and Dipeptide Repeat Proteins but Not Neurodegeneration in BAC Transgenic Mice. *Neuron.* 88(5):902-9.

Pieretti M, Zhang FP, Fu YH, Warren ST, Oostra BA, Caskey CT, Nelson DL. **(1991)** Absence of expression of the FMR-1 gene in fragile X syndrome. *Cell.* 66(4):817-22.

Pilli M, Arko-Mensah J, Ponpuak M, Roberts E, Master S, Mandell MA, Dupont N, Ornatowski W, Jiang S, Bradfute SB, Bruun JA, Hansen TE, Johansen T, Deretic V. **(2012)** TBK-1 promotes autophagy-mediated antimicrobial defense by controlling autophagosome maturation. *Immunity* 37(2):223-34.

Pottier C, Bieniek KF, Finch N, van de Vorst M, Baker M, Perkersen R, Brown P, Ravenscroft T, van Blitterswijk M, Nicholson AM, DeTure M, Knopman DS, Josephs KA, Parisi JE, Petersen RC, Boylan KB, Boeve BF, Graff-Radford NR, Veltman JA, Gilissen C, Murray ME, Dickson DW, Rademakers R. **(2015)** Whole-genome sequencing reveals important role for TBK1 and OPTN mutations in frontotemporal lobar degeneration without motor neuron disease. *Acta Neuropathol.* 130(1):77-92.

Pressman PS, Miller BL. **(2014)** Diagnosis and management of behavioral variant frontotemporal dementia. *Biol Psychiatry.* 75(7):574-81.

Primerano B, Tassone F, Hagerman RJ, Hagerman P, Amaldi F, Bagni C. **(2002)** Reduced FMR1 mRNA translation efficiency in fragile X patients with premutations. *RNA.* 8(12):1482-8.

Prudencio M, Belzil VV, Batra R, Ross CA, Gendron TF, Pregent LJ, Murray ME, Overstreet KK, Piazza-Johnston AE, Desaro P, Bieniek KF, DeTure M, Lee WC, Biendarra SM, Davis MD, Baker MC, Perkerson RB, van Blitterswijk M, Stetler CT, Rademakers R, Link CD, Dickson DW, Boylan KB, Li H, Petrucelli L. **(2015)** Distinct brain transcriptome profiles in C9orf72-associated and sporadic ALS. *Nat Neurosci.* 18(8):1175-82.

Phukan J, Pender NP, Hardiman O. **(2007)** Cognitive impairment in amyotrophic lateral sclerosis. *Lancet Neurol.* 6(11):994-1003.

Pulst SM, Nechiporuk A, Nechiporuk T, Gispert S, Chen XN, Lopes-Cendes I, Pearlman S, Starkman S, Orozco-Diaz G, Lunkes A, DeJong P, Rouleau GA, Auburger G, Korenberg JR, Figueroa C, Sahba S. **(1996)** Moderate expansion of a normally biallelic trinucleotide repeat in spinocerebellar ataxia type 2. *Nat Genet.* 14(3):269-76.

~R~

Renton AE, Majounie E, Waite A, Simón-Sánchez J, Rollinson S, Gibbs JR, Schymick JC, Laaksovirta H, van Swieten JC, Myllykangas L, Kalimo H, Paetau A, Abramzon Y, Remes AM, Kaganovich A, Scholz SW, Duckworth J, Ding J, Harmer DW, Hernandez DG, Johnson JO, Mok K, Ryten M, Trabzuni D, Guerreiro RJ, Orrell RW, Neal J, Murray A, Pearson J, Jansen IE, Sondervan D, Seelaar H, Blake D, Young K, Halliwell N, Callister JB, Toulson G, Richardson A, Gerhard A, Snowden J, Mann D, Neary D, Nalls MA, Peuralinna T, Jansson L, Isoviita VM, Kaivorinne AL, Hölttä-Vuori M, Ikonen E, Sulkava R, Benatar M, Wuu J, Chiò A, Restagno G, Borghero G, Sabatelli M; ITALSGEN Consortium, Heckerman D, Rogaeva E, Zinman L, Rothstein JD, Sendtner M, Drepper C, Eichler EE, Alkan C, Abdullaev Z, Pack SD, Dutra A, Pak E, Hardy J, Singleton A, Williams NM, Heutink P, Pickering-Brown S, Morris HR,

Ringholz GM, Appel SH, Bradshaw M, Cooke NA, Mosnik DM, Schulz PE. **(2005)** Prevalence and patterns of cognitive impairment in sporadic ALS. *Neurology.* 65(4):586-90.

Roberson ED, Hesse JH, Rose KD, Slama H, Johnson JK, Yaffe K, Forman MS, Miller CA, Trojanowski JQ, Kramer JH, Miller BL. **(2005)** Frontotemporal dementia progresses to death faster than Alzheimer disease. *Neurology.* 65(5):719-25.

Rogov V, Dötsch V, Johansen T, Kirkin V. **(2014)** Interactions between autophagy receptors and ubiquitin-like proteins form the molecular basis for selective autophagy. *Mol Cell.* 53(2):167-78.

Rosen DR, Siddique T, Patterson D, Figlewicz DA, Sapp P, Hentati A, Donaldson D, Goto J, O'Regan JP, Deng HX, et al. **(1993)** Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature.* 362(6415):59-62.

Ross OA, Rutherford NJ, Baker M, Soto-Ortolaza AI, Carrasquillo MM, DeJesus-Hernandez M, Adamson J, Li M, Volkening K, Finger E, Seeley WW, Hatanpaa KJ, Lomen-Hoerth C, Kertesz A, Bigio EH, Lippa C, Woodruff BK, Knopman DS, White CL 3rd, Van Gerpen JA, Meschia JF, Mackenzie IR, Boylan K, Boeve BF, Miller BL, Strong MJ, Uitti RJ, Younkin SG, Graff-Radford NR, Petersen RC, Wszolek ZK, Dickson DW, Rademakers R. **(2011)** Ataxin-2 repeat-length variation and neurodegeneration. *Hum Mol Genet.* 20(16):3207-12.

Rubinsztein DC. (2006) The roles of intracellular protein-degradation pathways in neurodegeneration. Nature. 443(7113):780-6.

Rubinsztein DC, Gestwicki JE, Murphy LO, Klionsky DJ. **(2007)** Potential therapeutic applications of autophagy.*Nat Rev Drug Discov.* 6(4):304-12.

Rutherford NJ, Zhang YJ, Baker M, Gass JM, Finch NA, Xu YF, Stewart H, Kelley BJ, Kuntz K, Crook RJ, Sreedharan J, Vance C, Sorenson E, Lippa C, Bigio EH, Geschwind DH, Knopman DS, Mitsumoto H, Petersen RC, Cashman NR, Hutton M, Shaw CE, Boylan KB, Boeve B, Graff-Radford NR, Wszolek ZK, Caselli RJ, Dickson DW, Mackenzie IR, Petrucelli L, Rademakers R. **(2008)** Novel mutations in TARDBP (TDP-43) in patients with familial amyotrophic lateral sclerosis. *PLoS Genet.* 4(9):e1000193.

Rutherford NJ, Heckman MG, Dejesus-Hernandez M, Baker MC, Soto-Ortolaza AI, Rayaprolu S, Stewart H, Finger E, Volkening K, Seeley WW, Hatanpaa KJ, Lomen-Hoerth C, Kertesz A, Bigio EH, Lippa C, Knopman DS, Kretzschmar HA, Neumann M, Caselli RJ, White CL 3rd, Mackenzie IR, Petersen RC, Strong MJ, Miller BL, Boeve BF, Uitti RJ, Boylan KB, Wszolek ZK, Graff-Radford NR, Dickson DW, Ross OA, Rademakers R. **(2012)** Length of normal alleles of C9ORF72 GGGGCC repeat do not influence disease phenotype. *Neurobiol Aging.* 33(12):2950.e5-7.

-S-

Sanpei K, Takano H, Igarashi S, Sato T, Oyake M, Sasaki H, Wakisaka A, Tashiro K, Ishida Y, Ikeuchi T, Koide R, Saito M, Sato A, Tanaka T, Hanyu S, Takiyama Y, Nishizawa M, Shimizu N, Nomura Y, Segawa M, Iwabuchi K, Eguchi I, Tanaka H, Takahashi H, Tsuji S. **(1996)** Identification of the spinocerebellar ataxia type 2 gene using a direct identification of repeat expansion and cloning technique, DIRECT. *Nat Genet.* 14(3):277-84.

Sareen D, O'Rourke JG, Meera P, Muhammad AK, Grant S, Simpkinson M, Bell S, Carmona S, Ornelas L, Sahabian A, Gendron T, Petrucelli L, Baughn M, Ravits J, Harms MB, Rigo F, Bennett CF, Otis TS, Svendsen CN, Baloh RH. **(2013)** Targeting RNA foci in iPSC-derived motor neurons from ALS patients with a C9ORF72 repeat expansion. *Sci Transl Med.* 5(208):208ra149.

Sasaki S (2011) Autophagy in spinal cord motor neurons in sporadic amyotrophic lateral sclerosis. *J Neuropathol Exp Neurol.* 70(5):349-59.

Sato K, Sugiyama T, Nagase T, Kitade Y, Ueda H. **(2014)** Threonine 680 phosphorylation of FLJ00018/PLEKHG2, a Rho family-specific guanine nucleotide exchange factor, by epidermal growth factor receptor signaling regulates cell morphology of Neuro-2a cells. *J Biol Chem.* 289(14):10045-56.

Satterfield TF, Pallanck LJ. **(2006)** Ataxin-2 and its Drosophila homolog, ATX2, physically assemble with polyribosomes. *Hum Mol Genet.* 15(16):2523-32.

Scotter EL, Vance C, Nishimura AL, Lee YB, Chen HJ, Urwin H, Sardone V, Mitchell JC, Rogelj B, Rubinsztein DC, Shaw CE. **(2014)** Differential roles of the ubiquitin proteasome system and autophagy in the clearance of soluble and aggregated TDP-43 species. *J Cell Sci.* 127(Pt 6):1263-78

Segura-Totten M, Wilson KL. **(2004)** BAF: roles in chromatin, nuclear structure and retrovirus integration. *Trends Cell Biol.* 14(5):261-6.

Sellier C, Rau F, Liu Y, Tassone F, Hukema RK, Gattoni R, Schneider A, Richard S, Willemsen R, Elliott DJ, Hagerman PJ, Charlet-Berguerand N. **(2010)** Sam68 sequestration and partial loss of function are associated with splicing alterations in FXTAS patients. *EMBO J.* 29(7):1248-61.

Sellier C, Freyermuth F, Tabet R, Tran T, He F, Ruffenach F, Alunni V, Moine H, Thibault C, Page A, Tassone F, Willemsen R, Disney MD, Hagerman PJ, Todd PK, Charlet-Berguerand N. **(2013)** Sequestration of DROSHA and DGCR8 by expanded CGG RNA repeats alters microRNA processing in fragile X-associated tremor/ataxia syndrome. *Cell Rep.* 3(3):869-80.

Shumaker DK, Lee KK, Tanhehco YC, Craigie R, Wilson KL. **(2001)** LAP2 binds to BAF.DNA complexes: requirement for the LEM domain and modulation by variable regions. *EMBO J.* 20(7):1754-64.

Sieben A, Van Langenhove T, Engelborghs S, Martin JJ, Boon P, Cras P, De Deyn PP, Santens P, Van Broeckhoven C, Cruts M. (2012) The genetics and neuropathology of frontotemporal lobar degeneration. *Acta Neuropathol.* 124(3):353-72.

Simonsen A, Tooze SA. **(2009)** Coordination of membrane events during autophagy by multiple class III PI3-kinase complexes. *J Cell Biol.* 186(6):773-82.

Skibinski G, Parkinson NJ, Brown JM, Chakrabarti L, Lloyd SL, Hummerich H, Nielsen JE, Hodges JR, Spillantini MG, Thusgaard T, Brandner S, Brun A, Rossor MN, Gade A, Johannsen P, Sørensen SA, Gydesen S, Fisher EM, Collinge J. **(2005)** Mutations in the endosomal ESCRTIII-complex subunit CHMP2B in frontotemporal dementia. *Nat Genet.* 37(8):806-8.

Slavoff SA, Mitchell AJ, Schwaid AG, Cabili MN, Ma J, Levin JZ, Karger AD, Budnik BA, Rinn JL, Saghatelian A. **(2013)** Peptidomic discovery of short open reading frame-encoded peptides in human cells. *Nat Chem Biol.* 9(1):59-64.

Snowden JS, Rollinson S, Thompson JC, Harris JM, Stopford CL, Richardson AM, Jones M, Gerhard A, Davidson YS, Robinson A, Gibbons L, Hu Q, DuPlessis D, Neary D, Mann DM, Pickering-Brown SM. **(2012)** Distinct clinical and pathological characteristics of frontotemporal dementia associated with C9ORF72 mutations. *Brain.* 135(Pt 3):693-708.

Sofola OA, Jin P, Qin Y, Duan R, Liu H, de Haro M, Nelson DL, Botas J. **(2007)** RNA-binding proteins hnRNP A2/B1 and CUGBP1 suppress fragile X CGG premutation repeat-induced neurodegeneration in a Drosophila model of FXTAS. *Neuron.* 55(4):565-71.

Sreedharan J, Blair IP, Tripathi VB, Hu X, Vance C, Rogelj B, Ackerley S, Durnall JC, Williams KL, Buratti E, Baralle F, de Belleroche J, Mitchell JD, Leigh PN, Al-Chalabi A, Miller CC, Nicholson G, Shaw CE. **(2008)** TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis. *Science.* 319(5870):1668-72.

Sudria-Lopez E, Koppers M, de Wit M, van der Meer C, Westeneng HJ, Zundel CA, Youssef SA, Harkema L, de Bruin A, Veldink JH, van den Berg LH, Pasterkamp RJ. **(2016)** Full ablation of C9orf72 in mice causes immune system-related pathology and neoplastic events but no motor neuron defects. *Acta Neuropathol.* 132(1):145-7.

Sullivan PM, Zhou X, Robins AM, Paushter DH, Kim D, Smolka MB, Hu F. **(2016)** The ALS/FTLD associated protein C9orf72 associates with SMCR8 and WDR41 to regulate the autophagy-lysosome pathway. *Acta Neuropathol Commun.* 4(1):51.

Suzuki N, Maroof AM, Merkle FT, Koszka K, Intoh A, Armstrong I, Moccia R, Davis-Dusenbery BN, Eggan K. **(2013)** The mouse C9ORF72 ortholog is enriched in neurons known to degenerate in ALS and FTD. *Nat Neurosci.* 16(12):1725-7.

Swash M, Desai J. **(2010)** Motor neuron disease: classification and nomenclature. *Amyotroph Lateral Scler Other Motor Neuron Disord.* 1(2):105-12.

Swinnen B, Robberecht W. **(2014)** The phenotypic variability of amyotrophic lateral sclerosis. *Nat Rev Neurol.* 10(11):661-70.

-T-

Tao Z, Wang H, Xia Q, Li K, Li K, Jiang X, Xu G, Wang G, Ying Z. **(2015)** Nucleolar stress and impaired stress granule formation contribute to C9orf72 RAN translation-induced cytotoxicity. *Hum Mol Genet.* 24(9):2426-41.

Tassone F, Hagerman RJ, Loesch DZ, Lachiewicz A, Taylor AK, Hagerman PJ. **(2000)** Fragile X males with unmethylated, full mutation trinucleotide repeat expansions have elevated levels of FMR1 messenger RNA.

Tassone F, Iwahashi C, Hagerman PJ. **(2004)** FMR1 RNA within the intranuclear inclusions of fragile X-associated tremor/ataxia syndrome (FXTAS). *RNA Biol.* 1(2):103-5.

Tassone F, Beilina A, Carosi C, Albertosi S, Bagni C, Li L, Glover K, Bentley D, Hagerman PJ. **(2007)** Elevated FMR1 mRNA in premutation carriers is due to increased transcription. *RNA.* 13(4):555-62.

Tassone F, Hagerman R. **(2012)** The fragile X-associated tremor ataxia syndrome. *Results Probl Cell Differ.* 54:337-57.

Teyssou E, Takeda T, Lebon V, Boillée S, Doukouré B, Bataillon G, Sazdovitch V, Cazeneuve C, Meininger V, LeGuern E, Salachas F, Seilhean D, Millecamps S. **(2013)** Mutations in SQSTM1 encoding p62 in amyotrophic lateral sclerosis: genetics and neuropathology. *Acta Neuropathol.* 2013 Apr;125(4):511-22. doi: 10.1007/s00401-013-1090-0.

Todd PK, Oh SY, Krans A, He F, Sellier C, Frazer M, Renoux AJ, Chen KC, Scaglione KM, Basrur V, Elenitoba-Johnson K, Vonsattel JP, Louis ED, Sutton MA, Taylor JP, Mills RE, Charlet-Berguerand N, Paulson HL. **(2013)** CGG repeat-associated translation mediates neurodegeneration in fragile X tremor ataxia syndrome. *Neuron.* 78(3):440-55.

Tran H, Almeida S, Moore J, Gendron TF, Chalasani U, Lu Y, Du X, Nickerson JA, Petrucelli L, Weng Z, Gao FB. **(2015)** Differential Toxicity of Nuclear RNA Foci versus Dipeptide Repeat Proteins in a Drosophila Model of C9ORF72 FTD/ALS. *Neuron.* 87(6):1207-14.

Traynor BJ, Codd MB, Corr B, Forde C, Frost E, Hardiman O. **(2000)** Amyotrophic lateral sclerosis mimic syndromes: a population-based study. *Arch Neurol.* 57(1):109-13.

Troakes C, Maekawa S, Wijesekera L, Rogelj B, Siklós L, Bell C, Smith B, Newhouse S, Vance C, Johnson L, Hortobágyi T, Shatunov A, Al-Chalabi A, Leigh N, Shaw CE, King A, Al-Sarraj S. **(2012)** An MND/ALS phenotype associated with C9orf72 repeat expansion: abundant p62-positive, TDP-43-negative inclusions in cerebral cortex, hippocampus and cerebellum but without associated cognitive decline. *Neuropathology.* 32(5):505-14.

Trottier Y, Lutz Y, Stevanin G, Imbert G, Devys D, Cancel G, Saudou F, Weber C, David G, Tora L, et al. **(1995)** Polyglutamine expansion as a pathological epitope in Huntington's disease and four dominant cerebellar ataxias. *Nature.* 378(6555):403-6.

U ~~

Urushitani M, Sato T, Bamba H, Hisa Y, Tooyama I. (2010) Synergistic effect between proteasome and autophagosome in the clearance of polyubiquitinated TDP-43. *J Neurosci Res.* 88(4):784-97.

-V-

van Blitterswijk M, Mullen B, Nicholson AM, Bieniek KF, Heckman MG, Baker MC, DeJesus-Hernandez M, Finch NA, Brown PH, Murray ME, Hsiung GY, Stewart H, Karydas AM, Finger E, Kertesz A, Bigio EH, Weintraub S, Mesulam M, Hatanpaa KJ, White CL 3rd, Strong MJ, Beach TG, Wszolek ZK, Lippa C, Caselli R, Petrucelli L, Josephs KA, Parisi JE, Knopman DS, Petersen RC, Mackenzie IR, Seeley WW, Grinberg LT, Miller BL, Boylan KB, Graff-Radford NR, Boeve BF, Dickson DW, Rademakers R. **(2014)** TMEM106B protects C9ORF72 expansion carriers against frontotemporal dementia. *Acta Neuropathol.* 127(3):397-406.

van Blitterswijk M, Gendron TF, Baker MC, DeJesus-Hernandez M, Finch NA, Brown PH, Daughrity LM, Murray ME, Heckman MG, Jiang J, Lagier-Tourenne C, Edbauer D, Cleveland DW, Josephs KA, Parisi JE, Knopman DS, Petersen RC, Petrucelli L, Boeve BF, Graff-Radford NR, Boylan KB, Dickson DW, Rademakers R. **(2015)** Novel clinical associations with specific C9ORF72 transcripts in patients with repeat expansions in C9ORF72. *Acta Neuropathol.* 130(6):863-76.

van Damme P, Veldink JH, van Blitterswijk M, Corveleyn A, van Vught PW, Thijs V, Dubois B, Matthijs G, van den Berg LH, Robberecht W. **(2011)** Expanded ATXN2 CAG repeat size in ALS identifies genetic overlap between ALS and SCA2. *Neurology.* 76(24):2066-72.

van Deerlin VM, Leverenz JB, Bekris LM, Bird TD, Yuan W, Elman LB, Clay D, Wood EM, Chen-Plotkin AS, Martinez-Lage M, Steinbart E, McCluskey L, Grossman M, Neumann M, Wu IL, Yang WS, Kalb R, Galasko DR, Montine TJ, Trojanowski JQ, Lee VM, Schellenberg GD, Yu CE. **(2008)** TARDBP mutations in amyotrophic lateral sclerosis with TDP-43 neuropathology: a genetic and histopathological analysis. *Lancet Neurol.* 7(5):409-16

van den Heuvel DM, Harschnitz O, van den Berg LH, Pasterkamp RJ. **(2014)** Taking a risk: a therapeutic focus on ataxin-2 in amyotrophic lateral sclerosis? *Trends Mol Med.* 20(1):25-35.

van Es MA, Veldink JH, Saris CG, Blauw HM, van Vught PW, Birve A, Lemmens R, Schelhaas HJ, Groen EJ, Huisman MH, van der Kooi AJ, de Visser M, Dahlberg C, Estrada K, Rivadeneira F, Hofman A, Zwarts MJ, van Doormaal PT, Rujescu D, Strengman E, Giegling I, Muglia P, Tomik B, Slowik A, Uitterlinden AG, Hendrich C, Waibel S, Meyer T, Ludolph AC, Glass JD, Purcell S, Cichon S, Nöthen MM, Wichmann HE, Schreiber S, Vermeulen SH, Kiemeney LA, Wokke JH, Cronin S, McLaughlin RL, Hardiman O, Fumoto K, Pasterkamp RJ, Meininger V, Melki J, Leigh PN, Shaw CE, Landers JE, Al-Chalabi A, Brown RH Jr, Robberecht W, Andersen PM, Ophoff RA, van den Berg LH. **(2009)** Genome-wide association study identifies 19p13.3 (UNC13A) and 9p21.2 as susceptibility loci for sporadic amyotrophic lateral sclerosis. *Nat Genet.* 41(10):1083-7.

van Langenhove T, van der Zee J, Van Broeckhoven C. **(2012)** The molecular basis of the frontotemporal lobar degeneration-amyotrophic lateral sclerosis spectrum. *Ann Med.* 44(8):817-28.

Vats A, Verma M, Gourie-Devi M, Taneja V **(2014)** C9ORF72 repeat expansion: a genetic mutation associated with Amyotrophic Lateral Sclerosis. *Current medicine research and practice.* 4 ; 161 -167.

Verkerk AJ, Pieretti M, Sutcliffe JS, Fu YH, Kuhl DP, Pizzuti A, Reiner O, Richards S, Victoria MF, Zhang FP, et al. **(1991)** Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome.
Cell. 65(5):905-14.

Vieira RT, Caixeta L, Machado S, Silva AC, Nardi AE, Arias-Carrión O, Carta MG. **(2013)** Epidemiology of early-onset dementia: a review of the literature. *Clin Pract Epidemiol Ment Health.* 9:88-95. doi: 10.2174/1745017901309010088.

-W-

Waite AJ, Bäumer D, East S, Neal J, Morris HR, Ansorge O, Blake DJ. **(2014)** Reduced C9orf72 protein levels in frontal cortex of amyotrophic lateral sclerosis and frontotemporal degeneration brain with the C9ORF72 hexanucleotide repeat expansion. *Neurobiol Aging.* 35(7):1779.e5-1779.e13.

Wang X, Xu S, Rivolta C, Li LY, Peng GH, Swain PK, Sung CH, Swaroop A, Berson EL, Dryja TP, Chen S. **(2002)** Barrier to autointegration factor interacts with the cone-rod homeobox and represses its transactivation function. *J Biol Chem.* 277(45):43288-300.

Wang X, Fan H, Ying Z, Li B, Wang H, Wang G. **(2010)** Degradation of TDP-43 and its pathogenic form by autophagy and the ubiquitin-proteasome system. *Neurosci Lett.* 469(1):112-6

Wang JY, Hessl DH, Hagerman RJ, Tassone F, Rivera SM. **(2012)** Age-dependent structural connectivity effects in fragile x premutation. *Arch Neurol.* 69(4):482-9.

Wang P, Miao CY. **(2012)** Autophagy in the disorders of central nervous system: vital and/or fatal? *CNS Neurosci Ther.* 18(12):955-6.

Webster CP, Smith EF, Bauer CS, Moller A, Hautbergue GM, Ferraiuolo L, Myszczynska MA, Higginbottom A, Walsh MJ, Whitworth AJ, Kaspar BK, Meyer K, Shaw PJ, Grierson AJ, De Vos KJ. **(2016)** The C9orf72 protein interacts with Rab1a and the ULK1 complex to regulate initiation of autophagy. *EMBO J.*;35(15):1656-76.

Wen X, Tan W, Westergard T, Krishnamurthy K, Markandaiah SS, Shi Y, Lin S, Shneider NA, Monaghan J, Pandey UB, Pasinelli P, Ichida JK, Trotti D. **(2014)** Antisense proline-arginine RAN dipeptides linked to C9ORF72-ALS/FTD form toxic nuclear aggregates that initiate in vitro and in vivo neuronal death. *Neuron.* 84(6):1213-25.

West RJ, Lu Y, Marie B, Gao FB, Sweeney ST. **(2015)** RAB8, POSH, and TAK1 regulate synaptic growth in a Drosophila model of frontotemporal dementia. *J Cell Biol.* 208(7):931-47.

Wheeler TM, Thornton CA. **(2007)** Myotonic dystrophy: RNA-mediated muscle disease. *Curr Opin Neurol.* 20(5):572-6.

Wijesekera LC, Leigh PN. **(2009)** Amyotrophic lateral sclerosis. *Orphanet J Rare Dis.* 4:3.

Woods SC. (2009) The control of food intake: behavioral versus molecular perspectives. *Cell Metab.*;9(6):489-98.

-X-

Xi Z, Zinman L, Moreno D, Schymick J, Liang Y, Sato C, Zheng Y, Ghani M, Dib S, Keith J, Robertson J, Rogaeva E. **(2013)** Hypermethylation of the CpG island near the G4C2 repeat in ALS with a C9orf72 expansion. *Am J Hum Genet.* 92(6):981-9.

Xi Z, van Blitterswijk M, Zhang M, McGoldrick P, McLean JR, Yunusova Y, Knock E, Moreno D, Sato C, McKeever PM, Schneider R, Keith J, Petrescu N, Fraser P, Tartaglia MC, Baker MC, Graff-Radford NR, Boylan KB, Dickson DW, Mackenzie IR, Rademakers R, Robertson J, Zinman L, Rogaeva E. **(2015)** Jump from pre-mutation to pathologic expansion in C9orf72. *Am J Hum Genet.* 96(6):962-70.

Xiao S, MacNair L, McGoldrick P, McKeever PM, McLean JR, Zhang M, Keith J, Zinman L, Rogaeva E, Robertson J. **(2015)** Isoform-specific antibodies reveal distinct subcellular localizations of C9orf72 in amyotrophic lateral sclerosis. *Ann Neurol.* 78(4):568-83.

Xiao S, MacNair L, McLean J, McGoldrick P, McKeever P, Soleimani S, Keith J, Zinman L, Rogaeva E, Robertson J. **(2016)** C9orf72 isoforms in Amyotrophic Lateral Sclerosis and Frontotemporal Lobar Degeneration. *Brain Res.* 1647:43-9.

Xie Z, Klionsky DJ. **(2007)** Autophagosome formation: core machinery and adaptations. Nat Cell Biol. 2007 Oct;9(10):1102-9.

-Y-

Yang M, Liang C, Swaminathan K, Herrlinger S, Lai F, Shiekhattar R, Chen JF. **(2016)** A C9ORF72/SMCR8-containing complex regulates ULK1 and plays a dual role in autophagy. *Sci Adv.* 2(9):e1601167.

Yokoseki A, Shiga A, Tan CF, Tagawa A, Kaneko H, Koyama A, Eguchi H, Tsujino A, Ikeuchi T, Kakita A, Okamoto K, Nishizawa M, Takahashi H, Onodera O. **(2008)** TDP-43 mutation in familial amyotrophic lateral sclerosis. *Ann Neurol.* 63(4):538-42.

-Z-

Zago S, Poletti B, Morelli C, Doretti A, Silani V. **(2011)** Amyotrophic lateral sclerosis and frontotemporal dementia (ALS-FTD). *Arch Ital Biol.* 149(1):39-56.

Zhang D, Iyer LM, He F, Aravind L. **(2012)** Discovery of Novel DENN Proteins: Implications for the Evolution of Eukaryotic Intracellular Membrane Structures and Human Disease. *Front Genet.* 3:283.

Zhang YJ, Jansen-West K, Xu YF, Gendron TF, Bieniek KF, Lin WL, Sasaguri H, Caulfield T, Hubbard J, Daughrity L, Chew J, Belzil VV, Prudencio M, Stankowski JN, Castanedes-Casey M, Whitelaw E, Ash PE, DeTure M, Rademakers R, Boylan KB, Dickson DW, Petrucelli L. **(2014)** Aggregation-prone c9FTD/ALS poly(GA) RAN-translated proteins cause neurotoxicity by inducing ER stress. *Acta Neuropathol.* 128(4):505-24.

Zheng R, Ghirlando R, Lee MS, Mizuuchi K, Krause M, Craigie R.**(2000)** Barrier-to-autointegration factor (BAF) bridges DNA in a discrete, higher-order nucleoprotein complex. *Proc Natl Acad Sci U S A.* 97(16):8997-9002.

Zu T, Gibbens B, Doty NS, Gomes-Pereira M, Huguet A, Stone MD, Margolis J, Peterson M, Markowski TW, Ingram MA, Nan Z, Forster C, Low WC, Schoser B, Somia NV, Clark HB, Schmechel S, Bitterman PB, Gourdon G, Swanson MS, Moseley M, Ranum LP. **(2011)** Non-ATG-initiated translation directed by microsatellite expansions. *Proc Natl Acad Sci U S A.* 108(1):260-5.

Zu T, Liu Y, Bañez-Coronel M, Reid T, Pletnikova O, Lewis J, Miller TM, Harms MB, Falchook AE, Subramony SH, Ostrow LW, Rothstein JD, Troncoso JC, Ranum LP. **(2013)** RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. *Proc Natl Acad Sci USA* 110(51):E4968-77.